MODEL SPECIFICATION SEARCHES IN LATENT GROWTH MODELING: A MONTE CARLO STUDY

A Dissertation

by

MINJUNG KIM

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

May 2012

Major Subject: Educational Psychology

Model Specification Searches in Latent Growth Modeling: A Monte Carlo Study

MODEL SPECIFICATION SEARCHES IN LATENT GROWTH MODELING:

A MONTE CARLO STUDY

A Dissertation

by

MINJUNG KIM

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Approved by:

Chair of Committee,	Oi-man Kwok
Committee Members,	Victor L. Willson
	Myeongsun Yoon
	Aaron Taylor
Head of Department,	Victor L. Willson

May 2012

Major Subject: Educational Psychology

ABSTRACT

Model Specification Searches in Latent Growth Modeling: A Monte Carlo Study.

(May 2012)

Minjung Kim, B.A., Hanyang University;

M.Ed., Texas A&M University

Chair of Advisory Committee: Dr. Oi-man Kwok

This dissertation investigated the optimal strategy for the model specification search in the latent growth modeling. Although developing an initial model based on the theory from prior research is favored, sometimes researchers may need to specify the starting model in the absence of theory. In this simulation study, the effectiveness of the start models in searching for the true population model was examined. The four possible start models adopted in this study were: the simplest mean and covariance structure model, the simplest mean and the most complex covariance structure model, the most complex mean and the simplest covariance structure model, and the most complex mean and covariance structure model. Six model selection criteria were used to determine the recovery of the true model: Likelihood ratio test (LRT), ΔCFI, ΔRMSEA, ΔSRMR, ΔAIC, and ΔBIC.

The results showed that specifying the most complex covariance structure (UN) with the most complex mean structure recovered the true mean trajectory most

successfully with the average hit rate above 90% using the ΔCFI, ΔBIC, ΔAIC, and ΔSRMR. In searching for the true covariance structure, LRT, ΔCFI, ΔAIC, and ΔBIC performed successfully regardless of the searching method with different start models.

DEDICATION

To my beautiful children, Aaron and Jason

ACKNOWLEDGEMENTS

First, I would like to say that I was so lucky to join this Research, Measurement, and Statistics program at Texas A&M University. I spent the most valuable time of my life in this program.

I wish to express my sincere gratitude to my advisor and committee chair, Dr. Oi-man Kwok. I would never have been able to finish my doctoral study without his support and encouragement. It was a great honor to have him as my advisor.

I would like to extend very special thanks to my committee members, Dr. Victor Willson, Dr. Myeongsun Yoon, and Dr. Aaron Taylor, for their guidance and support throughout the course of this research. I would also like to thank Dr. Deborah Simmons, Dr. Jorge Gonzalez, and Dr. Roger Goddard, who were my project advisors. As a quantitative methodologist, it was an invaluable experience to work with them for real data analyses.

A very special thanks to my best buddy, Eunju, who always stood by me. It would have been a lonely journey without her here. I would also like to thank other RMS folks (Catherine, Eun-sook, Jerry, Karen, Malena, Mark, Myung-hee, Susan, Russel, Yan, Yuan-yuan, and others) and my friends in South Korea (Jung-yeon & Jean) for their love and supports.

I would also like to say how much I appreciate my parents and parents-in-law for their endless support. Especially my mom and mother-in-law, they came here to take

care of my newborns (two times) for months so that I could move forward to finish my study. Thank you and love you, moms.

Finally, I want to truly thank my beloved partner, my husband, Sungmin. Without his help, I could not have made my dreams come true, a Ph.D. and two beautiful children.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

1. INTRODUCTION

Longitudinal data are common in educational and psychological studies. Many large scale longitudinal studies have been implemented in the last several decades, including: the National Education Longitudinal Study of 1988 (NELS-1998), the Early Childhood Longitudinal Study—Kindergarten Class of 1998-99 (ECLS-K), and the National Longitudinal Survey of Children and Youth (NLSCY). The two most commonly used approaches for analyzing longitudinal data are univariate analysis of variance (UANOVA) and multivariate analysis of variance (MANOVA). However, previous studies have shown limitations of these approaches, mainly due to their very restrictive assumptions. For example, the sphericity assumption for the UANOVA, which assumes a simple error covariance structure with constant variance for all repeated observations and constant covariance between all pairs of repeated observations, has been shown to be unrealistic for longitudinal data (Littell, Pendergast, & Natarajan, 2000). Although MANOVA has a relatively liberal assumption about the error covariance structure (i.e., placing no constraints on the error covariance structure), it can result in lower statistical power and convergence issues given that the error covariance structure is over parameterized (Everitt, 1998).

More recently developed approaches for analyzing longitudinal data include multilevel modeling (MLM) and latent growth modeling (LGM). Although these two approaches were developed under different frameworks, as Curran (2003) pointed out,

---

This dissertation follows the style of *Psychological Methods.*

they have been merging together with recent developments in computer programming. For example, most of the commercially developed structural equation modeling (SEM) programs, such as LISREL, EQS, and Mplus, have included a multilevel analysis routine. In brief, MLM (Goldstein, 1995; Hox, 2002; Raudenbush & Bryk, 2002) analyzes longitudinal data with an univariate format (i.e., each individual has multiple datalines to represent multiple observations), while LGM (Duncan, Duncan, Strycker, Li & Alpert, 1999; Preacher, Wichman, MacCallum, & Briggs, 2008) analyzes longitudinal data with a multivariate format (i.e., each time measure is represented by a variable and each individual only has a single dataline). The comparisons between the two approaches and the advantages of using the LGM over MLM have been discussed (e.g., MacCallum, Kim, Malarkey, & Kiecolt-Glaser, 1997; Kwok, Luo & West, 2010). One of the major advantages of using LGM over MLM is the availability of the overall model chi-square test along with a number of model fit indices for model evaluation.

Given that researchers do not know the true mean model in practice, it is important to search an optimal model by the model evaluation using those model fit indices. There were a few of studies on the sensitivity of fit indices to detect the model misspecification in LGM (Wu & West, 2010; Leite & Stapleton, 2011; Yu, 2002). Nevertheless, there have been few studies focusing primarily on model specification searches (i.e., searching for the optimal model) in LGM. Although it has been pointed out that the misspecification of the within-subject covariance structure had a substantial effect on the inference about growth parameters (Ferron, Dailey & Yi, 2002; Kwok, West & Green, 2007), the model specification search incorporating the error covariance

structure has been rarely investigated in the previous studies. In this document, I examine the model specification search strategy for building an optimal model in LGM. I start with a brief review of LGM, followed by a discussion of the impact of model misspecification. The specification search related literatures in SEM will be discussed last. I will summarize the current issues in searching for optimal LGM and propose some potential future research directions.

2. LITERATURE REVIEW OF MODEL SPECIFICATION SEARCH IN LATENT GROWTH MODELING

2.1 Latent Growth Modeling

Latent growth modeling (Duncan et al., 1999; Meredith & Tisak, 1990; Preacher et al., 2008), also known as latent growth curve modeling, is an advanced statistical methodology to model repeated measures by taking into account the inter-individual difference as well as the intra-individual difference under the structural equation modeling (SEM) framework. LGM can be viewed as a special case of SEM, which includes modeling both mean and covariance structures (Meredith & Tisak, 1990). In a linear latent growth model as presented in Figure 1, for example, the growth parameters (i.e., both intercept factor and linear growth/slope factor) are treated as unobserved/latent factors. Repeated measures are loaded on the two latent growth factors as multiple observed indicators. The intercept factor represents the level of the outcome measures at time equal to zero, which is sometimes treated as the measure of the initial status. The linear growth factor represents the linear rate of the change of the measure over time (Preacher et al., 2008). The linear latent growth model as presented in Figure 1 can be written as:

$$\underline{y} = \underline{\tau_y} + \Lambda_y \underline{\eta} + \underline{\varepsilon} \qquad [1]$$

where $\underline{y}$ (a 4 x 1 vector) refers to the observed outcome measure at each time point, $\underline{\tau_y}$ (a 4 x 1 vector) represents the intercepts of the measured variables, Λ is a 4 x 2 factor loading matrix showing the pattern of changes, $\underline{\eta}$ is a 2 x 1 vector containing the two

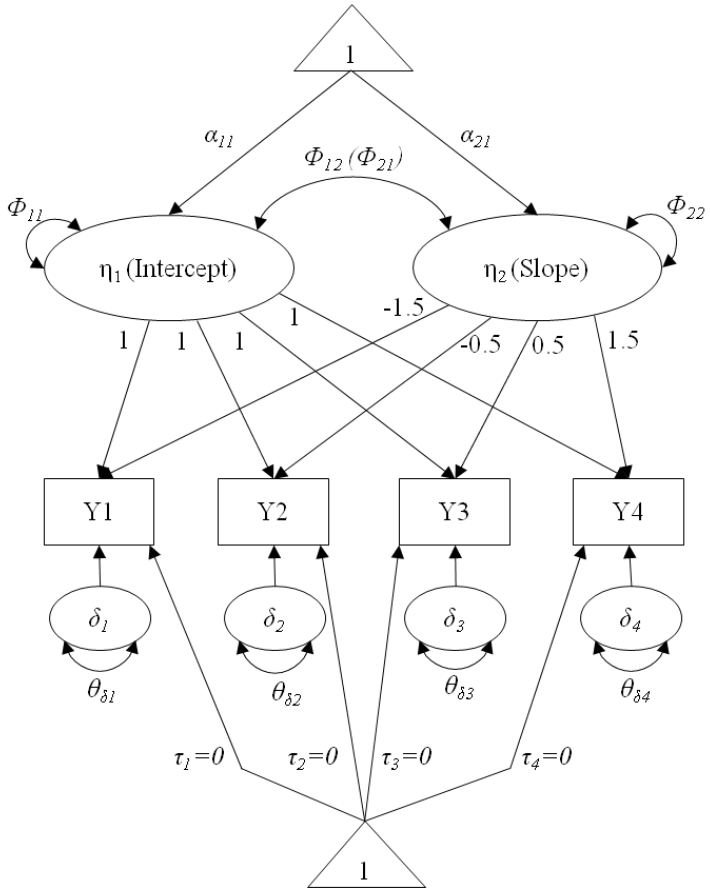

Figure 1. An example of the linear-trajectory latent growth modeling with four repeated measures having ID residual structure.

latent growth factors, and $\underline{\varepsilon}$ (a 4 x 1 vector) contains the corresponding error for each repeated measure. The two latent growth factors can be further written as:

$$\eta_{1i} = \alpha_{11} + \zeta_{1i} \quad [2]$$

$$\eta_{2i} = \alpha_{21} + \zeta_{2i} \quad [3]$$

where η_{1i} and η_{2i} are the latent intercept and slope factors, respectively; α_{11} refers to the mean of the intercept factor representing the level of the outcome measures at time equal to zero (or the initial time point); α_{21} refers to the mean of the slope factor representing

the linear growth rate of the outcome measures over time; and ζ_{1i} and ζ_{2i} represent the deviation between the corresponding individual parameter estimates from the mean estimates of the intercept and slope factors, respectively. The variance-covariance of equation (1) can be written as:

$$V\left(\underline{y}\right) = \Sigma = \Lambda_y \Psi \Lambda_y^{'} + \theta_\varepsilon \qquad [4]$$

where Ψ is a 2 x 2 matrix containing the variances and covariance of the two latent growth factors; Λ′ is the transpose of the Λ matrix which captures the overall pattern of change, and θ represents the matrix of variances and covariance among the errors (or unique factors). The expectation of equation (1) is:

$$E\left(\underline{y}\right) = \underline{\mu_y} = \underline{\tau_y} + \Lambda_y \underline{\alpha} \qquad [5]$$

where $\underline{\tau_y}$ is the latent intercept vector. The expectation of $\underline{y}$ is also known as the mean structure of the latent growth model. Given that both mean and covariance structures are the key for the estimation of LGM, I will elaborate on these structures again in the next section. Figure 1 shows the unconditional linear latent trajectory model of the four repeated measures. In multilevel models, the default residual variance structure for the outcome measures is the homogeneous variance model in which residual variances are constrained to be identical across the four time points. This default residual variance structure also has zero correlation between any pair of unique time points (i.e., identity structure). This is a very restrictive setting that can result in the reduction of statistical power (Ferron et al., 2002; Kwok et al., 2007). On the other hand, the error variance-covariance structure can be flexibly specified and estimated in LGM.

2.1.1 Mean structure in latent growth models

Traditional SEM models focus primarily on modeling the variances and covariances between the observed and latent variables without directly taking the mean structure into account. In other words, the mean structure, or the information related to the means of the variables, is not generally considered when models are estimated and evaluated (Kline, 1998). Mean structure, however, presents important information in longitudinal data. In the linear LGM example as presented in Figure 1, mean structure conveys information for means of the two growth related latent factors, which are the mean of intercept factor (i.e., α_{11}) and the mean of linear growth factor (i.e., α_{21}). In addition, the mean structure includes estimates the intercept of the repeated measured variables, which are typically constrained to be zero for model identification (Preacher et al., 2008). Means of latent factors are represented in the following mean matrix:

$$\alpha = \begin{bmatrix} \alpha_{11} \\ \alpha_{21} \end{bmatrix} \quad [6]$$

where α_{11} refers to the mean or average score of the outcome variable at the initial status (time=0) and α_{21} represents the average rate of change/growth over time. Intercepts of measured variables are represented in the following mean vector:

$$\tau_y = \begin{bmatrix} \tau_1 \\ \tau_2 \\ \tau_3 \\ \tau_4 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \quad [7]$$

in which all the intercepts are generally constrained to zero for model identification.

There are many functions of time transformation (e.g., piecewise, polynomial, exponential, and logarithmic) that can be used for growth models. Given that polynomial

functions are most commonly used in the social sciences (Ryoo, 2011), this study focused on model search using polynomial functions.

2.1.2 Residual variance-covariance structure in latent growth models

Residual variance-covariance structures convey the relationships among errors of the observed variables and latent variables. By recalling equation [4], the corresponding variance-covariance matrix of the linear growth model based on Figure 1 model 1 is shown below:

$$\Sigma = \Lambda_y \Psi \Lambda_y^{'} + \theta_\varepsilon = \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{bmatrix} \times \begin{bmatrix} \Phi_{11} & \Phi_{12} \\ \Phi_{21} & \Phi_{22} \end{bmatrix} \times \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & 3 \end{bmatrix} + \begin{bmatrix} \theta_\varepsilon & 0 & 0 & 0 \\ 0 & \theta_\varepsilon & 0 & 0 \\ 0 & 0 & \theta_\varepsilon & 0 \\ 0 & 0 & 0 & \theta_\varepsilon \end{bmatrix} \quad [8]$$

in which:

$$\Lambda_y = \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{bmatrix} \quad [9]$$

with the first column (all 1s) representing the average intercept estimate to be invariant across repeated measures while the second column represents the linear rate of change over time. The Φ matrix as shown in equation [10] below,

$$\Phi = \begin{bmatrix} \Phi_{11} & \Phi_{12} \\ \Phi_{21} & \Phi_{22} \end{bmatrix} \quad [10]$$

contains the variances and covariance of the two growth latent factors, with Φ_{11} representing the variance of the intercept factor and Φ_{22} representing the variance of the growth factor. Φ_{12} indicates the covariance (or magnitude of the covariation) between the two latent factors. All the parameter estimates in this matrix capture the

between/inter-individual variation/variance. On the other hand, the error variance-covariance matrix, θ_δ, has the following default structure under the MLM framework:

$$\theta_\varepsilon = \begin{bmatrix} \theta_\delta & 0 & 0 & 0 \\ 0 & \theta_\delta & 0 & 0 \\ 0 & 0 & \theta_\delta & 0 \\ 0 & 0 & 0 & \theta_\delta \end{bmatrix} \quad [11]$$

where $\theta_\delta$refers to the constant residual variance over time with uncorrelated residuals for each pair of time measures. This error variance-covariance captures the intra-individual difference across all the repeated measures. In addition to the default identity structure, there are some other commonly used error variance-covariance structures in longitudinal analysis. These additional covariance structures are important to consider for specifying the correct/optimal residual variance-covariance structure.

2.1.3 Types of variance-covariance structures

One of the strong advantages of using the LGM within the SEM framework to analyze longitudinal data is that it allows the residual structures (i.e., variance-covariance structures) to be freely specified based on the theory. There are several covariance structures typically used in longitudinal data analysis, including: identity structure, compound symmetry structure, first order autoregressive structure, Toeplitz structure, banded main diagonal structure, and the unstructured structure. Each of these error variance-covariance structures is briefly reviewed in the following sections.

2.1.3.1 Identity structure (ID)

The identity structure (ID) is the simplest variance-covariance structure, which assumes the identical variances across all the repeated measures without allowing covariance between any pair of the variables:

$$\mathrm{ID} = \begin{bmatrix} \sigma^2 & 0 & 0 & 0 \\ 0 & \sigma^2 & 0 & 0 \\ 0 & 0 & \sigma^2 & 0 \\ 0 & 0 & 0 & \sigma^2 \end{bmatrix} \quad [12].$$

This condition, however, is hardly ever met in longitudinal data since the repeated measures within an individual are typically correlated (Littell et al., 2000).

2.1.3.2 Compound symmetry structure(CS)

Compound symmetry (CS) assumes that within-individual variances are identical across the repeated measures, and the correlations between any two repeated variables are the same regardless of the distance between the time points. In other words, all of the diagonal elements are the same and all of the off-diagonal elements have the same value.

$$\mathrm{CS} = \begin{bmatrix} \sigma^2 + \sigma_1 & \sigma_1 & \sigma_1 & \sigma_1 \\ \sigma_1 & \sigma^2 + \sigma_1 & \sigma_1 & \sigma_1 \\ \sigma_1 & \sigma_1 & \sigma^2 + \sigma_1 & \sigma_1 \\ \sigma_1 & \sigma_1 & \sigma_1 & \sigma^2 + \sigma_1 \end{bmatrix} \quad [13].$$

2.1.3.3 First-Order Autoregressive Structure (AR (1))

One of the variance-covariance structures taking into account the distance of the repeated measure is the first-order autoregressive structure (AR(1)). The AR(1) structure specifies the equal variances across the repeated measures and auto-correlation ρ (i.e., $0 \leq \rho \leq 1$) for adjacent observations. The correlation between any pair of two time points is a function of ρ^d, where d refers to the distance between the two time points. For example, the first and second time point is ρ (i.e., $\rho^1 = \rho$), the first and third time point is ρ^2, and so on. The AR(1) structure is generally used when the measures are equally spaced.

$$\mathrm{AR}(1) = \sigma^2 \begin{bmatrix} 1 & \rho & \rho^2 & \rho^3 \\ \rho & 1 & \rho & \rho^2 \\ \rho^2 & \rho & 1 & \rho \\ \rho^3 & \rho^2 & \rho & 1 \end{bmatrix} \quad [14]$$

2.1.3.4 First-Order Autoregressive Moving Average Structure (ARMA(1,1))

The ARMA(1,1) model is similar to the AR(1) model to the extent that the correlation between a pair of two time points is a function of ρ^d . ARMA(1,1) includes an additional moving average parameter, γ, to the AR(1) structure. The ARMA(1,1) model with lag-1 process can be represented as $y_t = \rho_1 y_{t-1} + \varepsilon_t - \gamma_1 \varepsilon_{t-1}$. Like the AR(1) structure, the ARMA(1,1) structure has been commonly used for time series analysis.

$$\mathrm{ARMA}(1,1) = \sigma^2 \begin{bmatrix} 1 & \gamma & \gamma\rho & \gamma\rho^2 \\ \gamma & 1 & \gamma & \gamma\rho \\ \gamma\rho & \gamma & 1 & \gamma \\ \gamma\rho^2 & \gamma\rho & \gamma & 1 \end{bmatrix} \quad [15]$$

2.1.3.5 Toeplitz (TOEP) Structure

The Toeplitz structure (TOEP) can be viewed as a more general type of variance-covariance structure than the AR(1) structure. Unlike the AR(1) structure, which requires the power function of ρ between the time points, TOEP can specify any value between 0 and 1 for the correlation between any pair of time points laying on the same band. The TOEP structure is generally applied to longitudinal data with equal spacing measures.

$$\mathrm{TOEP} = \sigma^2 \begin{bmatrix} 1 & \rho_1 & \rho_2 & \rho_3 \\ \rho_1 & 1 & \rho_1 & \rho_2 \\ \rho_2 & \rho_1 & 1 & \rho_1 \\ \rho_3 & \rho_2 & \rho_1 & 1 \end{bmatrix} \quad [16]$$

2.1.3.6 Banded Main Diagonal Structure (UN(1))

The banded main diagonal (UN(1)) structure specifies a completely general (unstructured) covariance matrix for the first band of matrix, which is the variance components on the main diagonal. All the covariance components on the higher bands are set equal to zero.

$$\mathrm{UN}(1) = \begin{bmatrix} \sigma_1^2 & 0 & 0 & 0 \\ 0 & \sigma_2^2 & 0 & 0 \\ 0 & 0 & \sigma_3^2 & 0 \\ 0 & 0 & 0 & \sigma_4^2 \end{bmatrix} \quad [17]$$

2.1.3.7 Unstructured (UN) Structure

The most general variance-covariance structure is the unstructured structure (UN) in which each pair of time points can have its unique correlation.

$$\mathrm{UN}(1) = \begin{bmatrix} \sigma_1^2 & \sigma_{12} & \sigma_{13} & \sigma_{14} \\ \sigma_{21} & \sigma_2^2 & \sigma_{23} & \sigma_{24} \\ \sigma_{31} & \sigma_{32} & \sigma_3^2 & \sigma_{34} \\ \sigma_{41} & \sigma_{42} & \sigma_{43} & \sigma_4^2 \end{bmatrix} \quad [18]$$

All other variance-covariance matrices mentioned above are special cases of the UN structure. Since other matrices can be reproduced from the UN structure as constraining some parts of the variances and covariances, they are nested within the UN structure. Although specifying the UN structure may improve the model fit by relaxing more parameters, in many cases, it may come with a non-convergence issue, especially with more complex models along with small sample size.

2.2 Model Evaluation

In addition to the flexibility of modeling the error variance-covariance structure for longitudinal data, another advantage of using the LGM over MLM is the availability of the overall model chi-square test. Unlike MLM, LGM can also use a wide variety of model fit indices (e.g., CFI, RMSEA, SRMR, etc.). The overall model chi-square test compares the hypothesized model implied variance-covariance matrix with the input data matrix, while model fit indices show how similar or different the hypothesized model implied variance-covariance matrix and the input data variance-covariance matrix are. The overall model chi-square test and some of the model fit indices commonly used in the SEM, such as CFI, RMSEA and SRMR, are reviewed below.

2.2.1 Chi-square statistic

The χ^2 goodness-of-fit statistic estimates the amount of discrepancy between the model-implied variance-covariance matrices and the data-input variance-covariance matrices. With the maximum likelihood estimation method, the corresponding fit function is written as:

$$F_{ML} = \ln|\sum_0| + tr[S(\sum_0)^{-1}] - \ln|S| - r \qquad [19]$$

where S and $\sum_0$ refer to the observed variance-covariance matrix and model implied variance-covariance matrix, respectively; r refers to the number of observed variables in the model. The product of the maximum likelihood fitting function and the overall sample size minus one, approximately,

$$\chi^2_{ML} = (N-1)F_{ML} \qquad [20]$$

follows a central chi-square distribution if the null hypothesis is true. Degrees of freedom are equal to the difference between the number of unique pieces of information in the input data variance-covariance matrix and the number of freely estimated parameters in the model-implied variance-covariance matrix.

2.2.2 Comparative Fit Index (CFI)

The Comparative Fit Index (Bentler, 1995) is an incremental model fit index, which measures the degree of model fit improvement for the hypothesized model relative to a baseline model (i.e., independence null model). The CFI is ranged between zero and one; a value larger than one set to one; a value smaller than zero set to zero. A value closer to one refers to a good-fitting model; a value above .95 is good, between .95 and .90 is acceptable, and below .90 represents a poor-fitting model (Hu & Bentler, 1999). The complexity of the model is not considered in CFI. It is written as:

$$\mathrm{CFI} = 1 - \frac{\max\left[(\mathrm{T_T} - \mathrm{df_T}), 0\right]}{\max\left[(\mathrm{T_T} - \mathrm{df_T}), (\mathrm{T_B} - \mathrm{df_B}), 0\right]} \qquad [21]$$

where T_T and T_B refer to the T statistic value for the target model and baseline model, respectively, and df_T and df_B are the degrees of freedom for each corresponding model.

2.2.3 Root Mean Square of Error Approximation (RMSEA)

The Root Mean Square of Error Approximation (Steiger & Lind, 1980; Steiger, 1990) estimates the amount of discrepancy between the hypothesized model and the population model variance-covariance structure:

$$\mathrm{RMSEA} = \sqrt{\frac{\max\left[\left(\mathrm{T_T} - \mathrm{df_T}\right), 0\right]}{\mathrm{df_T}(\mathrm{N} - 1)}} \qquad [22]$$

where T_T refers to the T statistic for the target model, df_T is the degrees of freedom, and N is the sample size. As shown in the formula, RMSEA penalizes models with a large number of free parameters (small degrees of freedom). In other words, all else being equal, a complex model has a larger RMSEA value than a simple model. A value below .05 refers to a good model, between .05 and .08 is acceptable, and above .10 indicates a poor-fitting model (Browne & Cudeck, 1993).

2.2.4 Standardized Root Mean Square Residual (SRMR)

The Standardized Root Mean Square Residual (Bentler, 1995; Jöreskog & Sörbom, 1981) is another measure of the standardized average discrepancy between the model implied variance-covariance matrix and the input data variance-covariance matrix:

$$\text{SRMR} = \sqrt{\left\{2\sum_{i=1}^{P},\sum_{j=1}^{i}\left[\frac{s_{ij}-\hat{\sigma}_{ij}}{s_{ii}s_{jj}}\right]^2\right\}/p(p+1)} \qquad [23]$$

where s_{ij} refers to the observed covariances from the input data variance-covariance matrix, $\hat{\sigma}_{ij}$ is the reproduced covariances from the model implied variance-covariance matrix, s_{ii} and s_{jj} are the observed standard deviations, and p is the number of observed variables. The cutoff criteria of SRMR are similar to those values of RMSEA; that is, a value less than .05 indicates a good fit, a model with a value between .05 and .08 is acceptable, and a value greater than .10 indicates a poor fit (Hu & Bentler, 1999). SRMR is an absolute fit index, which does not take into account the model complexity.

2.3 Model Misspecifications in Latent Growth Models

Although the specification of the residual variance-covariance structure in general has limited impact on the estimation of the mean trajectory, it has a considerable

effect on the estimation of the standard error of the mean trajectory parameters and even the overall model evaluation (Ferron et al., 2002; Kwok et al., 2007; Sivo, Fan, & Witta, 2005). Hence, it is important to specify the correct/optimal residual variance-covariance structure. For repeated measures, identity structure (ID) is commonly used as a default in most MLM statistical software (e.g., HLM, SPSS Mixed routine, SAS proc mixed routine, and STATA xtmixed routine). As pointed out by Kwok and colleagues (2007), utilizing the MLM framework, under-specifying the error variance-covariance structure can result in an over-estimated standard error, which in turn, can reduce the statistical power for testing the average growth trajectory parameter.

Grimm and Widaman (2010) examined two common residual structures (i.e., identity structure and unstructured) in LGM and proposed two alternatives for specifying the residual structure from the idea regarding growth curve reliability. According to Grimm and Widaman, growth curve reliability is based on the ratio of the intercept variance to the total variance. As shown in their study, researchers should examine and apply alternative residual structures rather than using the default error variance-covariance structure from the statistical software. Grimm and Widaman (2010) further introduced the growth curve reliability as one of the residual structures to be considered for analyzing longitudinal data.

In LGM, both mean and variance-covariance structures can have substantial effect on the overall model chi-square test and other fit indices (Wu, West, & Taylor, 2009). Nonetheless, few studies have examined the sensitivity of the model fit indices on detecting misspecification in LGM. Wu and West (2010) examined the sensitivity of the

model fit indices on detecting misspecifications in both mean and variance-covariance structure, and they suggested using a saturated covariance structure to capture the misspecifications in the mean structure. For example, when a linear growth model (hypothesized model) is specified for a dataset which has a true quadratic growth pattern, the model misspecification in the mean structure was detected by estimating all the elements in the covariance structure. Nevertheless, Grimm and Widaman (2010) found none of the model fit indices they examined were consistently sensitive to all kinds of model misspecifications. For example, SRMR was more sensitive to misspecifications in the covariance structure, while RMSEA and CFI were more sensitive to misspecifications in the interaction between mean and covariance structures. In addition, further studies on model specification search, which utilizes their recommendation (i.e., saturating the covariance structure to capture the misspecification of the mean structure) are needed.

2.4 Specification Searches

2.4.1 Model Specification Search

Specification search (Leamer & Leamer, 1978; Long, 1983) was termed as a modification of the prior model to improve the model fit considering its parsimony (MacCallum, 1986). Model specification searches have been studied in terms of the SEM framework and covariance structure modeling (MacCallum, 1986; MacCallum, Roznowski, & Necowitz, 1992). According to MacCallum (1986), model specification searches are typically conducted with the following four steps:

i) Specify a model representing the relationship among the measured variables and latent variables based on the theory of the research field

ii) Fit the model to sample data

iii) Assess the model fit and parameter estimates

iv) Modify the model to improve the model fit considering model parsimony.

When step iii) shows poor model fit, specification searches are necessary to find the meaningful model not only showing adequate model fit but also having its parsimony leading to interpretability. The poorly fitted model can be improved by model comparison using the model fit indices. Unlike MLM, most SEM computer software (e.g., Mplus, LISREL, EQS, Amos, etc.) provide the model fit indices, such as, CFI, RMSEA, SRMR, and so on. Given that the model specification search is an exploratory procedure, researchers are required to not only build a statistical model but also improve the model to fit the data. By using those model fit indices, researchers can proceed through the model specification search with consideration of the start model and the parameter addition or deletion method (Peugh & Enders, 2010).

2.4.2 Starting Model

Specifying the theory-based initial model will be the most reasonable way to start the specification search. For example, if the researchers expect to observe a linearly growing pattern, they would specify the linear growth model; if they expect to observe a linearly increasing and then decreasing pattern, they might choose the quadratic growth model. Although developing an initial model based on the theory from prior research is favored, sometimes researchers may need to specify the starting model in the absence of

theory. An alternative approach is starting from either the null model (the simplest model) or the saturated model (the most complex model). The null model in the LGM is also called the "intercept-only" model leaving only an intercept factor, deleting the growth factor from the model (Widaman & Thompson, 2003). Figure 2 shows the intercept-only model with constrained equal residual variances across time. On the other hand, the saturated model (see Figure 3) is defined as a model with "… as many parameter estimates as can be made from the data, a number of estimates that equals the number of unique elements of the covariance matrix and mean vector…" (Widaman & Thompson, 2003). Sometimes a more complex form of the growth model, such as quadratic growth model or piece-wise growth model, can be used as the saturated model. The null model and the saturated model are thought of being on a continuum of nested models. Widaman and Thompson (2003) presented the continuum of nested models graphically as

$$M_0 \ldots\ldots\ldots M_k \ldots\ldots\ldots M_l \ldots\ldots\ldots M_s,$$

where M_0 refers to the null model and M_s refers to the saturated model; M_0 is nested within M_k, M_k is nested within M_l, and M_l is nested within M_s. In terms of the variance-covariance structure, ID structure can be considered as the null model with the most restricted constraints, while UN structure can be regarded as the saturated model with the fewest constraints. In this study, the performance of finding the true model with four different starting points (i.e., (1) ID covariance structure for intercept-only model, (2) UN covariance structure for intercept-only model, (3) ID covariance structure for the highest polynomial order growth model, (4) UN covariance structure for the highest

polynomial order growth model) will be investigated. The four possible starting points finding a true model (e.g., quadratic growth with AR(1) structure) are presented in Figure 4.

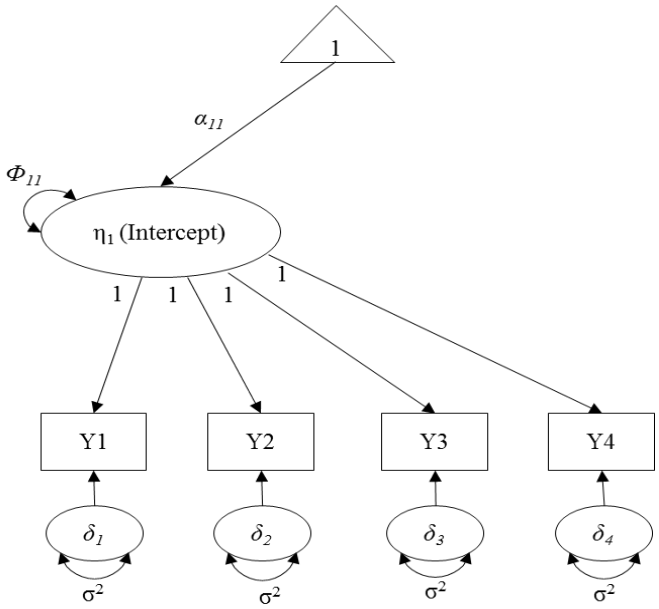

Figure 2. An example of intercept-only model with identity (ID) variance-covariance structure.

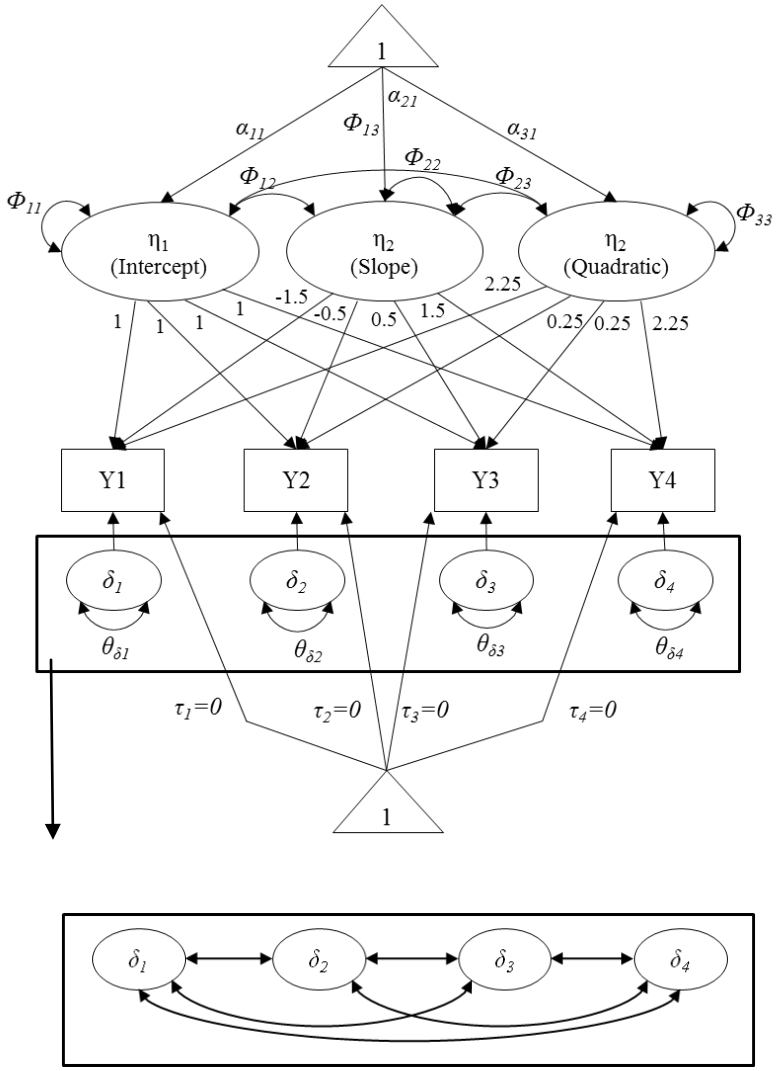

Figure 3. An example of quadratic growth model with saturated (UN) variance-covariance structure

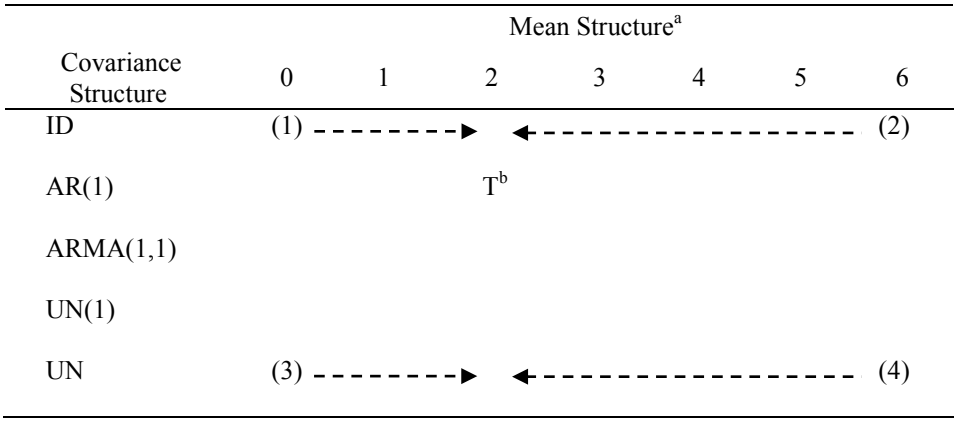

Figure 4. Four starting models in searching for the population mean structure ([a]Degree of polynomial order growth terms (i.e., 0=intercept-only, 1=linear growth, 2=quadratic growth 3=cubic growth, 4=4th order, 5=5th order, and 6=6th order polynomial model); [b]T=True model for Study 1 Scenario 1

2.4.3 Parameter addition/deletion method

There are two approaches for conducting the specification searches to find the true population model: parameter addition and deletion methods. With the parameter addition method, a null model (M_0) is generally chosen as a starting model. First, the null model is tested and the chi-square statistic is calculated based on the simplest model. Modification indices can be used for identifying which parameter needs to be freely estimated for improving the model fit. After relaxing the constraint, the modified model is tested and the corresponding chi-square statistic is obtained. The chi-square difference tests are conducted to examine the significance of the model fit improvement. On the other hand, in the parameter deletion method, a saturated model (the most complicated model) is chosen as the starting model and other simpler alternative models are then estimated. The final model is determined by the chi-square difference tests. Most of the previous studies regarding the model search have shown that the parameter deletion

method (also known as the backward elimination technique) performed better than the parameter addition method in recovering the population model given that it includes all possible significant coefficients in the model (Green, Thompson, & Poirier, 1999; Peugh & Enders, 2010; Suyapa, Silvia, & MacCallum, 1988).

On the other hand, Ryoo (2011) examined the performance of "step-up" (or parameter addition) and "top-down" (parameter deletion) model building approaches on searching for the true growth model under the linear mixed model (LMM) framework. Ryoo used the intercept-only model as the null starting model for the step-up search while he used the highest order polynomial term (e.g., 9^{th} order polynomial term for 10 wave data) as the saturated starting model for the top-down search. In his simulation study, the fixed effects, which are associated with the growth pattern and the static predictors, and the random effects (i.e., between-subject variance-covariance structure) were manipulated for the model specification search. However, the random error structure (i.e. within-subject covariance structure) was not considered as selecting the optimal model but fixed to an ID structure for all conditions. Although the study showed that the step-up approach performed better than the top-down method, the effect of specifying the random error structure needs to be looked into. As with Ryoo's study, in terms of the error structure, the simplest ID structure model (i.e., forcing variances to be equal across the repeated time points) is typically used as a default starting model in most of the multilevel modeling software packages (Ferrer, Hamagami, & McArdle, 2004). There is no study specifically examining the effectiveness of these two

approaches in searching for the optimal mean and variance-covariance structures in LGM, and further research on this is needed.

2.4.4 Model comparison

There are two types of approaches for comparing two competing models in terms of the nested structure. Nested structure means that one model can be reproduced from the other model by either imposing or relaxing constraint(s). For example, a linear growth model with ID variance-covariance structure can be compared to a linear growth model with UN variance-covariance structure given the same data; the more restricted model (the ID variance-covariance structure) can reproduce the UN structure by freely estimating all the elements in the variance-covariance matrix. Nested models can be compared by using the likelihood ratio testss. When the models to be compared are not nested, the change in the goodness of fit indices (ΔGFI; Cheung & Rensvold, 2002) and information criteria, such as Akaike's Information Criteria (AIC) and Bayesian Information Criteria (BIC), can be used for model comparison as well as the LRT. These model comparison strategies are further reviewed as following.

2.4.4.1 The Likelihood Ratio Chi-Square Statistics

The likelihood ratio test (LRT) is used for model comparison between nested models. The LRT is also known as the chi-square difference test (Bollen, 1989). The chi-square difference ($\Delta\chi^2$) is calculated as

$$\Delta\chi^2 = \chi^2{}_0 - \chi^2{}_1 \qquad [24]$$

where $\chi^2{}_0$ is the value for the simpler/reduced model and $\chi^2{}_1$ refers to the value for the alternative (liberal) model. This difference between the model chi-square statistics ($\Delta\chi^2$)

follows the chi-square distribution with degrees of freedom equal to the difference of the degrees of freedom between the two models (i.e., $\Delta df = df_0 - df_1$).if the null hypothesis is true. Significant $\Delta\chi^2$ (e.g., $\Delta\chi^2(1) > 3.84$) results in rejecting the null-hypothesis and concluding that the more liberal model fits the data better. To the extent that the chi-square difference test is sensitive to the sample size (i.e., more likely to reject the null hypothesis with large sample size), there have been suggestions for using alternative fit indices for model comparison (Cheung & Rensvold, 2002; Chen, 2007).

2.4.4.2 Alternative goodness-of-fit indices

Due to the plausible drawbacks of using the chi-square difference test, alternative goodness-of-fit indexes (GFI) have been developed for model comparison. Cheung and Rensvold (2002) investigated the change of GFIs (i.e., ΔGFIs) for testing the measurement invariance. They examined the performance of twenty ΔGFIs based on different model fit indices (e.g., ΔCFI, ΔRMSEA, etc.) to detect the invariance constraints when testing measurement invariance. ΔGFI is defined as,

$$\Delta GFI = GFI_c - GFI_{uc} \quad [25]$$

where GFI_c and GFI_{uc} refer to the GFI value of constrained (simpler) and unconstrained (more complex) models, respectively. For example, ΔCFI was calculated by $CFI_c - CFI_{uc}$. The result showed that many ΔGFI performed better than $\Delta\chi^2$ because they are more robust to the small sample size relative to $\Delta\chi^2$. Among the twenty ΔGFIs, ΔCFI, ΔGamma hat, and ΔMcDonald's Noncentrality Index performed well for testing the measurement invariance under the confirmatory factor analysis framework. For example, Cheung and Rensvold (2002) suggested that the absolute value of ΔCFI smaller than or

equal to 0.01 should not be considered as significant between the constrained and unconstrained models even though the chi-square difference between the two models may be significant. Chen (2007) recommended using ΔCFI supplemented by ΔRMSEA and ΔSRMR with cutoff values of 0.01, 0.015, and 0.01, respectively, for the measurement invariance testing.

Nevertheless, as pointed out by Fan and Sivo (2009), these ΔGFIs are purely descriptive and should not be used for statistical inference, such as claiming that one model is better than the other one. Moreover, in their simulation study, Fan and Sivo (2009) also showed that ΔGFIs were not appropriate for mean structure invariance analysis because ΔGFIs were sensitive to the model size (i.e., the complexity of the model). They suggested using the traditional $\Delta\chi^2$ tests rather than using ΔGFIs for testing invariance, especially when testing invariance related to the mean structure.

2.4.4.3 Akaike Information Criterion (AIC)

The Akaike Information Criterion (AIC) was developed by Akaike (1974). AIC can be computed as:

$$\mathrm{AIC} = -2\log L + 2p \qquad [26]$$

where logL refers to the maximized log-likelihood value and p is the number of parameters to be estimated (Weakliem, 2004). AIC penalizes the number of parameters to be estimated; that is, AIC gets bigger when the model gets more complicated. AIC can be viewed as an index of badness of fit. Nevertheless, the value of AIC carries no particular meaning (i.e., it is non-interpretable) by itself. In general, AIC is used for model comparison, and models with smaller AIC are more preferable. The major

advantage of using AIC (and other information criteria) over the LRT is that the competing models need not to be nested. As long as the same data were used for estimating the models, the AIC can be used as an index for model selection.

2.4.4.4 Bayesian Information Criterion (BIC)

The Bayesian Information Criterion (BIC) is also known as Schwarz Criterion (Schwarz, 1978). BIC can be written as follows:

$$\text{BIC} = -2\log L + p\log(N) \qquad [27]$$

where logL is the maximized log-likelihood value, p is the number of parameters, and N is the number of cases. Similar to AIC, BIC also penalizes complex models, and a smaller BIC value indicates a more favorable model when comparing models. As Weakliem (2004) pointed out, BIC differs from AIC in terms of the inclusion of sample size (N) and the size of the penalty (i.e., $2p$ and $p\log(N)$).

Nevertheless, the results from previous simulation studies did not show much support of the effectiveness of using these information criteria for model selection. For example, Keselman, Algina, Kowalchuk, and Wolfinger (1998) showed that the hit rates of AIC and BIC (i.e., the rate of correctly identifying the true variance-covariance structure in longitudinal analysis) were only 47% and 35%, respectively.

There were some guidelines for interpreting the absolute difference of the information criteria (i.e., Δinformation criteria) between the two competing models. For AIC, Burnham and Anderson (1998) suggested using at least 4 as the definite evidence to determine that one model (with smaller AIC) is statistically better than the other model (with larger AIC). Likewise, Raftery (1996) argued that the absolute difference of

BIC between the two compared models should be at least 2 to indicate the definite difference. Following these guidelines, I used the cutoff value of 4 for AIC and 2 for BIC difference. Details of how to use these Δinformation criteria for model search is illustrated in the method section.

3. STUDY 1: MODEL SPECIFICATION SEARCH FOR THE OPTIMAL MEAN STRUCTURE IN LATENT GROWTH MODELING (LGM): A MONTE CARLO STUDY

Longitudinal data are common in educational and psychological studies. Many large scale longitudinal studies have been implemented in the last several decades. More recently developed approaches for analyzing those longitudinal data include multilevel modeling (MLM) and latent growth modeling (LGM). Nevertheless, there have been few studies focusing primarily on model specification search in LGM.

While traditional structural equation modeling (SEM) takes only the variance-covariance structure into account, LGM considers both mean and variance-covariance structures for establishing a model. Curran, Obeidat, and Losardo (2010) pointed out that one of the most important steps in building any growth model is to identify the optimal functional form of the trajectory over time. However, only a few studies have been conducted regarding the model misspecification issues in both mean and covariance structure in LGM. Wu and West (2010) have studied the misspecification in both mean and residual structures in LGM and suggested saturating either the mean or the variance-covariance structure to identify possible misspecification in the other structure. This approach is opposite to the common practice, especially in MLM in which the simplest structure (e.g., ID structure for the error variance-covariance matrix) is used as the default when analyzing longitudinal data (Ferrer et al., 2004; Grimm & Widaman, 2010).

The issue of how to obtain the optimal growth model (in terms of both mean and variance-covariance structures) is not completely clear yet.

Because of the issue presented above, the major goal of this simulation study is to investigate the effectiveness of start models and model selection criteria in searching for the optimal mean and variance-covariance structures in LGM. Given that researchers are in general more interested in obtaining the correct mean structure (i.e., the average growth trajectory), the search sequence in our simulation study started with finding the correct mean trajectory, followed by the search for the correct covariance structure. The starting model contained either the simplest mean structure (e.g., intercept only model) or the most complex structure (e.g., the highest possible polynomial growth structure), along with either the simplest covariance structure (i.e., one invariant residual variance) or the most complex covariance structure (i.e., unequal variances and covariances over repeated times). In this study, with respect to the between-subject covariance structure (i.e., growth parameter covariance structures), the UN structure (i.e., estimating all the elements in Φ matrix) was used as the default. Therefore, the performance of the four possible start models (Figure 4) were investigated to determine how successfully they recover the true growth model: (1) the simplest covariance and mean structures, (2) the simplest covariance with the most complex mean structure, (3) the most complex covariance with the simplest mean structure, and (4) the most complex covariance and mean structures.

This study focused on examining the performance of searching the population model with different starting points using several model selection criteria—LRT, ΔGFI

(i.e., ΔCFI, ΔRMSEA, and ΔSRMR), and information criteria (i.e., AIC and BIC)—that have been recommended in previous simulation studies (Chen, 2007; Cheung & Rensvold, 2002; Raftery, 1996; Burnham & Anderson, 1998).

1. What start model and what model comparison strategies should be used in searching for the optimal mean growth trajectory?
    a) Which start model performs best to recover the population mean trajectory among (1) simplest covariance with simplest mean structure, (2) simplest covariance with most complex/saturated mean structure, (3) most complex/saturated covariance with simplest mean structure, and (4) most complex/saturated covariance and mean structure?
    b) What model fit indices or model comparison strategies perform best to find the optimal mean structure among LRT, information criteria, and ΔGFI?
2. What start model and what model comparison strategies should be used in searching for the optimal within-subject covariance/residual structure?
    a) Which starting covariance structure (i.e., the simplest ID structure or the saturated UN structure) performs best to recover the population error structure with specifying the correct mean structure?

Which model comparison strategy can result in the highest hit rate (i.e., correctly identifying the true covariance structure model)?

I considered several simulation conditions: different sample size conditions, number of measurement waves, and population mean and residual structures. The ideal search method should be able to find the true/target model regardless of the simulation

conditions. The details of how I generated the target model and evaluated the performance of four start models and the searching tools (e.g., model fit indices and information criteria) are presented below. In this study, my focus was on examining the effectiveness of different start models and different model selection indices/criteria to search the optimal mean structure for repeated measures.

3.1 Scenario 1: true linear growth model

In the first scenario, the true/target model was a linear growth model with the AR(1) residual structure.

3.1.1 Method for scenario 1

The simulation used a 2 (number of measurement waves: 4 or 8) × 2 (number of individuals: 210 or 390) factorial design to generate the data. A total of 2,000 replications were generated for each condition using the Mplus6.11 Monte Carlo procedure (Muthén & Muthén, 1998-2011), yielding 8,000 total datasets. All data were generated under Mplus with a multivariate normal distribution. Each dataset was then analyzed using the four start models separately along with different types of model selection indices/criteria. The details of each design factor are described below together with a justification of the values selected for the study.

3.1.1.1 Number of participants and measurement waves

According to Khoo, West, Wu, and Kwok (2006), the mean number of individuals in the multiwave longitudinal studies published in *Developmental Psychology* for the past couple of decades was 210 (SD=180). In this simulation study, 210 individuals were used for the condition with small number of participants, and 390

individuals (i.e., 1SD above the mean number of participants) were used for the condition with large number of participants. Additionally, I chose 4 waves as the small number of repeated measures and 8 waves as the medium number of measures based on the same review by Khoo and colleagues (2006). To reduce the potential multicollinearity problem (Biesanz, Deeb-Sossa, Papadakis, Bollen, & Curran, 2004), I centered the time variable to the mean so zero represented the middle point of the waves (i.e., $Time_{4waves}$ = [−1.5, −.5, .5, 1.5], and $Time_{8waves}$ = [−3.5, −2.5, −1.5, −.5, .5, 1.5, 2.5, 3.5]).

3.1.1.2 Mean growth trajectory

An average linear growth trajectory was used as the true/target model, and the population parameters for this growth trajectory were selected based on prior simulation studies on longitudinal data analysis (Kwok et al., 2007). The following effect size equation (Raudenbush & Liu, 2001) was used to select the population parameters:

$$\delta = \frac{\alpha}{\tau_{11}} \qquad [28],$$

where δ is the standardized effect size, α is the average growth parameter, and τ_{11} is the variance of the random effect associated with the growth parameter. For the average growth trend, medium effect size (δ = 0.5; Cohen, 1988) was adopted to generate the population data. In a similar manner, medium size of variances of growth parameters (τ_{11} = 0.10; Raudenbush & Liu, 2001) was selected. Given the values of δ and τ_{11}, the corresponding average linear growth rate, α_{21},was computed to be 0.16. The average intercept, α_{11}, was fixed to a constant value in all conditions (i.e., α_{11} = 0.10) leading to $\alpha = \begin{bmatrix} 0.10 \\ 0.16 \end{bmatrix}$.

3.1.1.3 Type of variance-covariance structure

The AR(1) structure was used for generating the within-subject covariance structure given that it was commonly used in time series analysis (Velicer & Fava, 2003;West & Hepworth, 1991). The residual variances of the measurement waves (i.e., θ_δ) were set as 1.00, which is a common practice in power analysis (Bosker, Snijders, & Guldemond, 2003; Snijders & Bosker, 1993). Based on prior simulation studies on error structures (Hamaker, Dolan, & Molenaar, 2002; Sivo & Willson, 2000), the autocorrelation coefficient, ρ, was set to 0.80.

3.1.1.4 Magnitude of phi matrix

The phi matrix (equation 10) represents the variances and covariance(s) of the growth latent factors; this matrix captures the between-individual variation on growth trajectories. According to Raudenbush and Liu (2001), the medium size of slope variance (Φ_{22}) is 0.10. To the extent that intercept variance has generally been larger than the variation of the growth trends in longitudinal studies, Φ_{11} was set to 0.20 constantly for all conditions. Therefore, for the linear growth model, the variances and covariance of the latent factors were set to be $\Phi=\begin{bmatrix}.20 & .05\\ .05 & .10\end{bmatrix}$, where Φ_{11} was the intercept variance and Φ_{22} was the slope variance. Since the covariance between intercept and slope has generally been smaller than the variances, it was set to 0.05, which was half of the slope variance.

3.1.1.5 Start models and specification search method

For model specification search, I considered the four start models as presented in Figure 4: (1) simplest covariance and mean structures, (2) simplest covariance and most

complex mean structures, (3) most complex covariance and simplest mean structures, and (4) most complex covariance and mean structures. For 4-wave data, the most complex mean structure was a quadratic growth model, while it was a 6^{th} order polynomial model for 8-wave data. In practice, this saturated mean structure is not recommended because it is more likely to have a convergence problem due to having too many parameters in a model. The highest order polynomial model, however, was used as the most complex mean model in this simulation study to test the statistical significance.

Parameter addition/deletion methods were also considered for the specification search procedure along with the start models. As Ryoo (2011) named in his study, the parameter addition and deletion approaches related to the mean structure were called as a "step-up" and a "top-down" method in this study. For the step-up method, the simplest mean structure model (i.e., intercept-only model) was compared to the next simplest mean model (i.e., a linear growth model) with the same residual structure. If the result was statistically significant, a growth term was added to the model one at a time to examine the significance of the corresponding changing growth pattern. Searching was stopped when the result of the model comparison was not statistically significant, and a model with fewer parameters was selected as the best fitting mean structure.

In a similar manner, for the top-down method, the most complex mean structure starting model (e.g., 6^{th} order polynomial model for 8-wave data) was compared to a model with the next largest number of mean structure parameters (e.g., 5^{th} order polynomial model). If the result was statistically significant, searching was stopped, and a model with a higher order polynomial term was selected as the best fitting mean

structure. Searching proceeded as deleting a growth term one at a time if the comparison between the two adjacent mean models was not statistically significant. Searching was stopped when the comparison was statistically significant. All the searching procedures were conducted separately by each within-subject covariance structure.

For setting up the simplest covariance structure, only one element in the within-subject covariance structure (i.e., invariant residual variance across repeated measures) was freely estimated. As a default, the UN structure for the between-subject variance-covariance structure (i.e., estimating all the elements related to the growth parameters) was used. For the most complex covariance structure, the UN structure was selected as estimating all the elements in the within-subject covariance structure. Since all of the degrees of freedom were used to estimate the parameters associated with the within-subject covariance structure, elements in the between-subject covariance structure (i.e., variances and covariances of the growth parameters) were set to zero for model identification.

3.1.1.6 Evaluation criteria and specification search procedure

I evaluated three types of model selection criteria to search for the optimal mean structure: (1) LRT, (2) ΔGFI (i.e., ΔCFI, ΔRMSEA, and ΔSRMR), and (3) information criteria (i.e., ΔAIC & ΔBIC). Using those evaluation criteria, both step-up and top-down methods were examined. Given that a polynomial model with lower order term (e.g., intercept-only) was nested within a polynomial model with higher order term (e.g., linear model), LRT was used for model comparison. Regardless of nesting structure, ΔGFI (i.e., $GFI_c - GFI_{uc}$) and Δinformation criteria (i.e., ΔAIC and ΔBIC) were also used for

the model searching procedure. Following the guidelines for how to use the ΔGFI (Chen, 2007), searching was stopped with a value of less than 0.01 for ΔCFI, 0.015 for ΔRMSEA, and 0.01 for ΔSRMR. When the absolute difference of the information criteria between the two competing models was less than 4 for AIC (Burnham & Anderson, 1998) and 2 for BIC (Raftery,1996), searching was stopped. For example, if the ΔAIC was at least 4, the model comparison was statistically significant and searching was stopped. A model with more parameters (i.e., a model with more complex mean structure) was selected for the best fitting mean trajectory.

3.1.1.7 Dependent variable

The primary dependent variable was a hit rate of whether the true mean model (i.e., a linear growth model) was successfully searched by model selection criteria with different start models. For this dependent variable, correct model recovery was coded as a binary variable (i.e., 0 for fail and 1 for success) for all replicates by all conditions. The hit rate (i.e., percentage of replicates reaching the true mean structure) was summarized to report the performance of different searching tools.

3.1.2 Results for scenario 1

In the first scenario, I focused on searching for the optimal mean trajectory for the repeated measures. The performance of reaching the true mean structure from the four possible start models combined with a parameter addition/deletion method was evaluated by several model selection criteria. In the first scenario, the true population model was a linear growth model with the AR(1) residual structure. There were four

conditions with a combination of two measurement waves (i.e., 4 or 8) and two sample sizes (i.e., 210 or 390).

3.1.2.1 Convergence rate

Before conducting the model search using different start models, each dataset was analyzed with the correct model specification. Table 1 presents the convergence rate of the correctly specified model for four simulation conditions (i.e., two sample sizes and two measurement waves). Given that the convergence rate ranged from 54.10% to 76.95% by the simulation condition, the number of replications was increased to 4,000. The first 2,000 datasets of the converged replications with no errors were used for further model search study for all simulation conditions.

Table 2 indicates the convergence rate for the analyses with specifying different start models for the simulated data. All the analyses for 4-wave data were successfully converged. For 8-wave data, however, the results were not converged for the 4th-, 5th-,

Table 1. Percentage of convergence for specifying the true linear growth model for study 1 scenario 1

Mean[a]	Cov[b]	wave	*n*	Final #of successful replications[c]	%of replications
Linear	AR(1)	4	210	1082	54.10%
Linear	AR(1)	4	390	1166	58.30%
Linear	AR(1)	8	210	1418	70.90%
Linear	AR(1)	8	390	1539	76.95%

Note. [a]Mean: true linear growth model; [b]Cov: covariance structure was AR(1); [c]Number of replications = 2,000

Table 2. Percentage of convergence for specifying the various mean and covariance structure model for scenario 1

Wave	True mean	True cov[a]	Cov[b]	n	Mean[c]						
					0	1	2	3	4	5	6
4wave	Linear	AR(1)	id	210	100.00	100.00	100.00	-	-	-	-
4wave	Linear	AR(1)	id	390	100.00	100.00	100.00	-	-	-	-
4wave	Linear	AR(1)	un	210	100.00	100.00	100.00	-	-	-	-
4wave	Linear	AR(1)	un	390	100.00	100.00	100.00	-	-	-	-
8wave	Linear	AR(1)	id	210	100.00	100.00	100.00	100.00	74.55	75.35	74.25
8wave	Linear	AR(1)	id	390	100.00	100.00	100.00	100.00	80.30	82.50	82.15
8wave	Linear	AR(1)	un	210	100.00	100.00	100.00	100.00	100.00	100.00	99.95
8wave	Linear	AR(1)	un	390	100.00	100.00	100.00	100.00	100.00	100.00	99.90

Note. [a]True cov: the true covariance structure was AR(1); [b]Cov: specified covariance structure (i.e., id=Identity, un=Unstructured); [c]Mean: the degree of polynomial growth term (e.g., 0 = intercept-only model, 1 = linear growth model, 2 = quadratic growth model, etc.).

and 6th-order polynomial growth model when specifying the simplest ID residual structure. Given that the true model was a linear growth trajectory, there were too many parameter components being estimated that should have been fixed to zero. To reduce the problem of nonconvergence issue, I constrained the between-subject variance-covariance elements for the higher order polynomial terms (i.e., 4th, 5th, and 6th) to zero. After that, the convergence rates for the small sample size were 74.55%, 75.35%, and 74.25% for 4th-, 5th-, and 6th-order polynomial models, respectively. For the large sample size, the convergence rates were increased to 80.30%, 82.5%, and 82.15%, respectively.

3.1.2.2 True model recovery

Table 3 presents the hit rate of the model selection criteria. The first column indicates the starting point of the model search procedure in terms of the residual structures (see Figure 4). The second column identifies the starting point in terms of the covariance structure (i.e., ID = the simplest point, UN = the most complex point). In the next column, step-up refers to starting with the simplest mean structure, while top-down refers to starting with the most complex mean structure. Given that the linear model was located in the middle of the possible mean structure with 4-wave data (i.e., between the intercept-only and quadratic growth model), a step-up or a top-down method was not considered in this condition. For 8-wave data, a step-up search used the intercept-only model as a start model, while the top-down method started from the 6th order polynomial model. The sample size and number of measurement waves were also presented in the table. The next six columns report the number and percentage of replications that select

Table 3. True mean structure recovery rate (percentage of recovered data) by the model selection criteria[a]

Start[b]	Cov[c]	Mean[d]	*n*	Waves	LRT	ΔCFI	ΔRMSEA	ΔSRMR	ΔAIC	ΔBIC
-	ID	-	210	4	11.85	19.40	19.50	22.45	24.00	82.95
-	ID	-	390	4	0.15	4.85	8.70	11.35	1.35	32.50
(1)	ID	Step-up	210	8	0.00	0.00	0.00	7.35	0.00	0.00
(1)	ID	Step-up	390	8	0.00	0.00	0.00	1.75	0.00	0.00
(2)	ID	Top-down	210	8	0.00	0.00	0.00	0.45	0.00	0.00
(2)	ID	Top-down	390	8	0.00	0.00	0.00	0.15	0.00	0.00
-	UN	-	210	4	94.40	98.05	83.85	86.45	98.05	98.60
-	UN	-	390	4	94.35	99.85	84.80	93.55	98.30	99.35
(3)	UN	Step-up	210	8	95.25	93.90	92.15	94.85	98.80	99.40
(3)	UN	Step-up	390	8	95.10	99.00	92.90	97.95	98.15	99.40
(4)	UN	Top-down	210	8	77.10	93.90	52.30	91.70	92.45	96.60
(4)	UN	Top-down	390	8	77.05	99.00	58.75	97.50	92.60	97.25

Note. [a]True model: linear growth model with the AR(1) error structure; [b]Start: start model: (1) the simplest covariance and mean structure, (2) the simplest covariance and the most complex mean structure, (3) the most complex covariance and the simplest mean structure, (4) the most complex covariance and mean structure; [c]Cov: specification of the covariance structure (i.e., ID = identity, UN = unstructured); [d]Mean: Step-up = parameter addition method starting from an intercept-only model (i.e., intercept-only model →linear model), Top-down = parameter deletion method starting from the most complex model (i.e., 6th order polynomial model →5th order polynomial model →4th order polynomial model →cubic model → quadratic model → linear model).

the true mean model (i.e., linear model) as the best fitting mean structure by LRT and other model selection criteria.

Table 4 shows the details of the model search procedure, which provides the result of model comparisons between the two adjacent mean structures. In Table 4, the fourth column indicates the degree of the polynomial term for the two mean structures (i.e., 0 = intercept-only model, 1 = linear model, 2 = quadratic model, and so on). For example, the first row (i.e., ID covariance structure, wave=4, *n*=210, mean: 0 vs. 1) means that all replications (100.00%) favored the linear growth model over the intercept-only model by all model selection criteria.

Simplest residual structure. The first two rows in Table 3 present the correct model recovery rate for 4-wave data when the ID covariance structure was specified. As shown in the table, the overall hit rate of the model selection indices was higher with the small sample size (i.e., 30.03%) than with the larger size (i.e., 9.82%). For the sample size, the LRT hit rate was 11.85% and the average ΔGFI (i.e., ΔCFI, ΔRMSEA, and ΔSRMR) hit rate was 20.45%. However, the hit rates were reduced to 0.15% and 8.3% with large sample size. Likewise, ΔAIC performed better with the small sample size (24.00%) compared to the large sample size (1.35%). ΔBIC performed relatively well with the hit rate of 82.95% for the small sample size. As Table 4 presents, for ΔBIC, 17.05% of replications determined that the quadratic growth models were better fitting models than the linear growth (i.e., true mean) models. However, ΔBIC hit rate was also decreased to 32.50% for the large sample size as other comparison tools. Figure 5

Table 4. Percentage of replications favoring a model with more parameters[a]

Cov[b]	Wave	*n*	Mean[c]	LRT	ΔCFI	ΔRMSEA	ΔSRMR	ΔAIC	ΔBIC
ID	4	210	0 vs. 1	100.00	100.00	100.00	100.00	100.00	100.00
ID	4	210	1 vs. 2	88.15	80.6	80.5	77.55	76.00	17.05
ID	4	390	0 vs. 1	100.00	100.00	100.00	100.00	100.00	100.00
ID	4	390	1 vs. 2	99.85	95.15	91.3	88.65	98.65	67.50
ID	8	210	0 vs. 1	100.00	100.00	100.00	100.00	100.00	100.00
ID	8	210	1 vs. 2	100.00	100.00	100.00	92.65	100.00	100.00
ID	8	210	2 vs. 3	100.00	99.65	89.75	57.95	100.00	98.80
ID	8	210	3 vs. 4	74.35	68.95	61.60	12.65	73.85	52.30
ID	8	210	4 vs. 5	5.05	0.00	0.15	0.00	1.25	0.75
ID	8	210	5 vs. 6	3.30	0.00	0.00	0.00	0.80	0.30
ID	8	390	0 vs. 1	100.00	100.00	100.00	100.00	100.00	100.00
ID	8	390	1 vs. 2	100.00	100.00	100.00	98.25	100.00	100.00
ID	8	390	2 vs. 3	100.00	100.00	95.15	65.10	100.00	100.00
ID	8	390	3 vs. 4	80.30	78.95	75.60	11.55	80.30	79.55
ID	8	390	4 vs. 5	5.30	0.00	0.05	0.00	2.00	1.15
ID	8	390	5 vs. 6	3.75	0.00	0.00	0.00	0.90	0.35
UN	4	210	0 vs. 1	99.95	98.65	99.95	100.00	99.75	99.35
UN	4	210	1 vs. 2	5.55	0.60	16.10	13.55	1.70	0.75
UN	4	390	0 vs. 1	100.00	99.90	100.00	100.00	100.00	100.00
UN	4	390	1 vs. 2	5.65	0.05	15.20	6.45	1.70	0.65
UN	8	210	0 vs. 1	100.00	93.90	100.00	100.00	100.00	99.90
UN	8	210	1 vs. 2	4.75	0.00	7.85	5.15	1.20	0.50
UN	8	210	2 vs. 3	5.55	0.00	10.20	1.95	1.65	0.60
UN	8	210	3 vs. 4	6.00	0.00	11.45	0.70	2.15	0.95
UN	8	210	4 vs. 5	4.05	0.00	13.45	0.15	1.35	0.60
UN	8	210	5 vs. 6	4.95	0.00	16.50	0.35	1.45	0.65
UN	8	390	0 vs. 1	100.00	99.00	100.00	100.00	100.00	100.00
UN	8	390	1 vs. 2	4.90	0.00	7.10	2.05	1.85	0.60
UN	8	390	2 vs. 3	5.15	0.00	7.15	0.30	1.35	0.45
UN	8	390	3 vs. 4	5.45	0.00	9.50	0.15	1.25	0.65
UN	8	390	4 vs. 5	5.15	0.00	10.35	0.00	1.90	0.65
UN	8	390	5 vs. 6	5.10	0.00	14.50	0.00	1.30	0.40

Note. [a]True model: linear growth model with the AR(1) error structure; [b]Cov: specification of the covariance structure (i.e., ID = identity, UN = unstructured); [c]Mean: the degree of polynomial growth term. (e.g., 0 = intercept-only model, 1 = linear growth model, 2 = quadratic growth model, etc.)

presents the hit rates of the model selection indices when specifying the simplest covariance structure for 4-wave data.

When the number of measurement waves was increased to 8-wave, specifying the simplest residual structure performed poorly to recover the true mean structure regardless of the search method (step-up or top-down). Except ΔSRMR, all other fit indices or tests failed to search the true mean model with a recovery rate of zero. Although ΔSRMR with the step-up approach recovered 7.35% of true mean model for the small sample size condition, the rate went down to 1.75% with the large number of subjects. Figure 6 and Figure 7 report the performance of the model selection tools by the step-up and top-down method, respectively, when the simplest within-subject

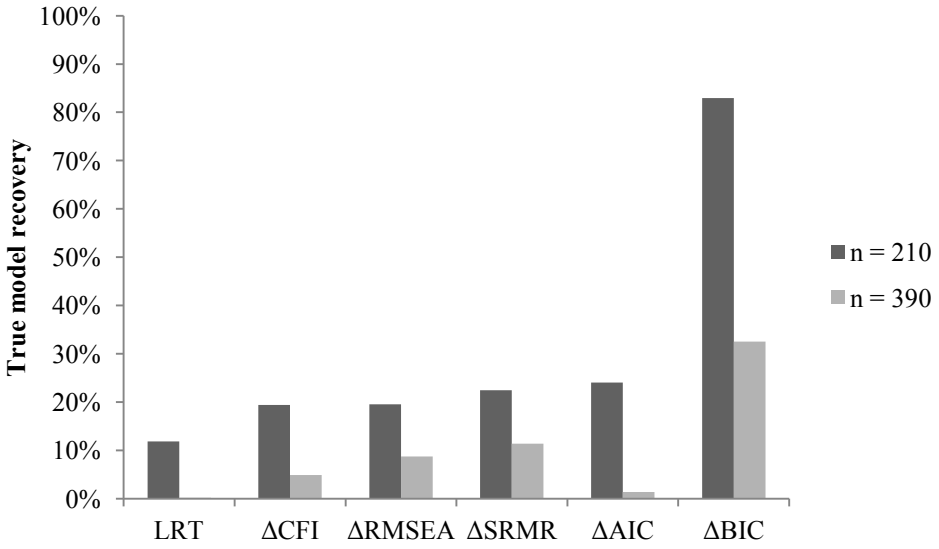

Figure 5. Hit rates of model selection criteria by specifying the simplest covariance structure
Note. True model = linear growth model with AR(1) structure for 4-wave data)

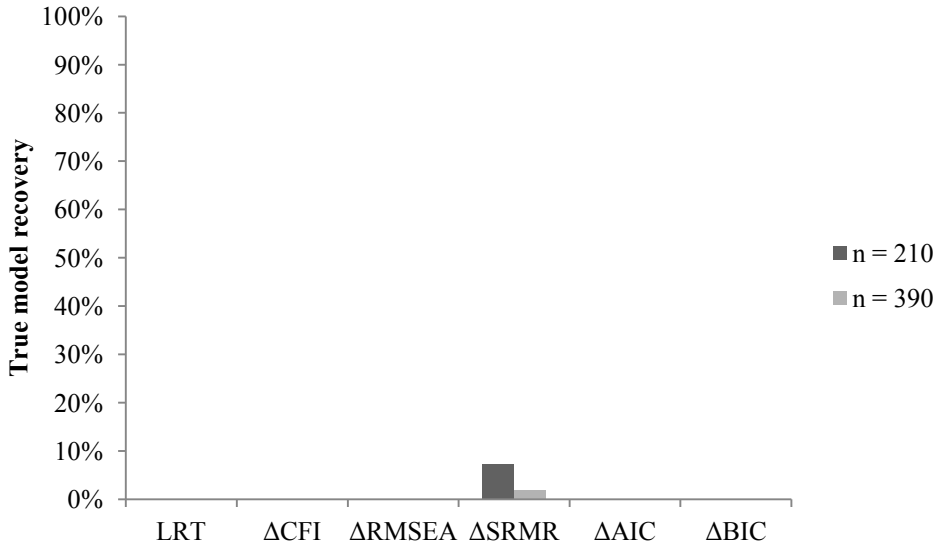

Figure 6. Hit rates of model selection criteria by specifying the simplest mean and simplest covariance structure
Note. True model = linear growth model with AR(1) structure for 8-wave data

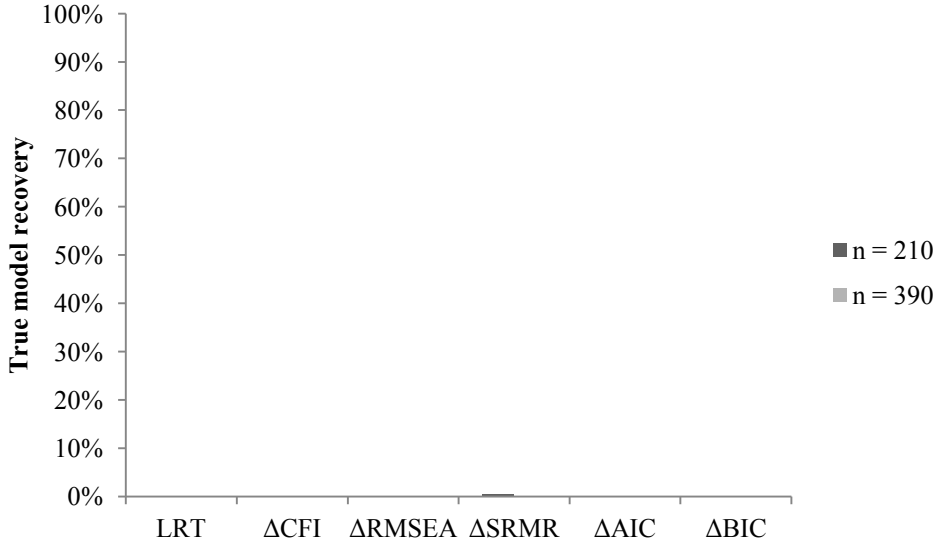

Figure 7. Hit rates of model selection criteria by specifying the most complex mean and simplest covariance structure
Note. True model = linear growth model with AR(1) structure for 8-wave data

covariance structure (ID) was specified. As shown in the figures, the average hit rates were substantially low, which was close to zero.

Most complex residual structure. On the other hand, when the most complex covariance structure (UN) was specified, all model selection criteria were able to search the optimal mean structure successfully regardless of the various simulation conditions. The average recovery rate of the model selection indices for 4-wave data was at or above 90% except the ΔRMSEA, which showed 84.33% of recovery. Figure 8 indicates the hit rate of the model selection indices for 4-wave data when specifying the UN structure.

For 8-wave data, the average hit rates of the two sample size condition for all model fit indices were above 90%, ranged from 92.53% (ΔRMSEA) to 99.40% (ΔBIC). The hit rates were higher with using the step-up method than with the top-down method except for the ΔCFI, which showed the same recovery rate on both methods. Figure 9 and Figure 10 present the hit rates of the model selection indices by the step-up and top-down method, respectively, with specifying the most complex residual structure (UN) for 8-wave data. The different performance between the simplest covariance structure model and the most complex covariance structure model is noticeable when observing the figures.

3.2 Scenario 2: true quadratic growth model

In the first scenario, the true linear growth trajectory was successfully searched by saturating the within-subject covariance structure. In the second scenario, I conducted additional simulations to replicate and extend the findings of the first study. The true population model was extended to include more types of within-subject covariance

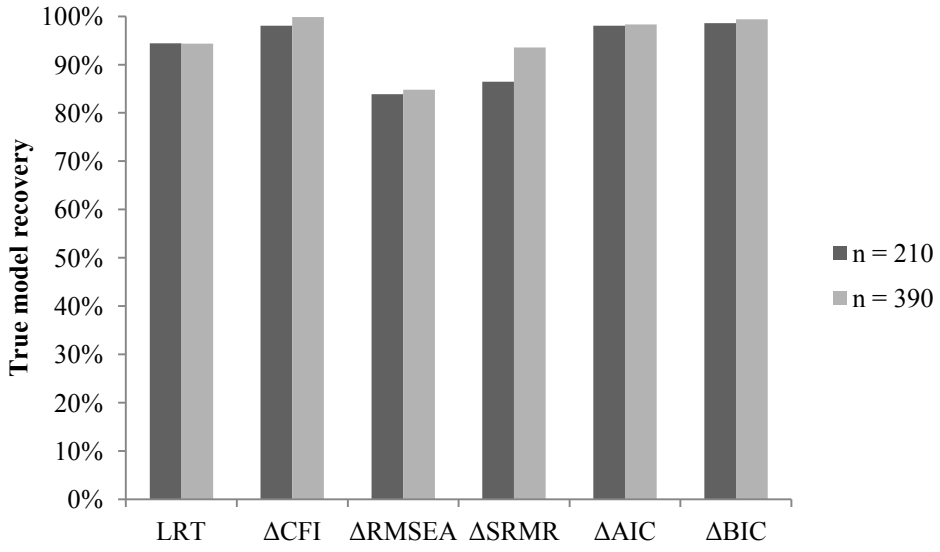

Figure 8. Hit rates of model selection criteria by specifying the saturated covariance structure

Note. True model = linear growth model with AR(1) structure for 4-wave data

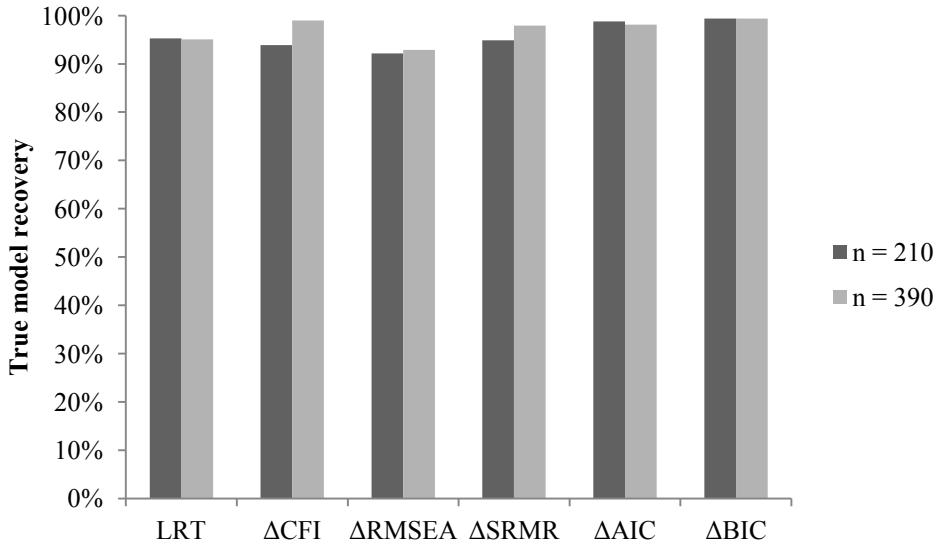

Figure 9. Hit rates of model selection criteria by specifying the simplest mean and saturated covariance structure

Note. True model = linear growth model with AR(1) structure for 8-wave data

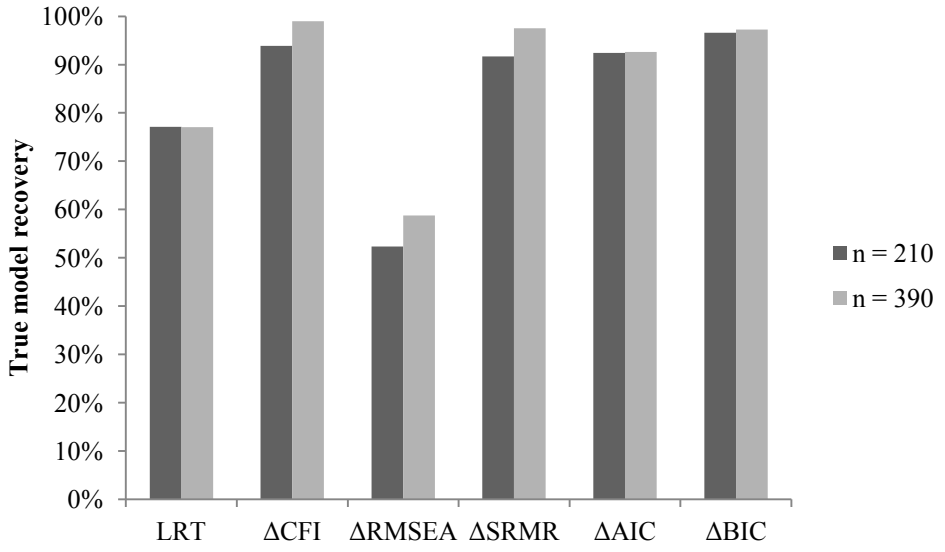

Figure 10. Hit rates of model selection criteria by specifying the most complex mean and saturated covariance structure
Note. True model = linear growth model with AR(1) structure for 8-wave data

structures (i.e., AR(1), UN(1), and ID). In this study, the true models were quadratic growth models with the saturated between-subject covariance structure (i.e., non-zero variances and covariances of the growth latent factors).

3.2.1 Method for scenario 2

Most simulation conditions were identical with the first scenario but differed on the number of waves and residual structures. Given that the true model was a quadratic growth model with increased number of parameters, only 8-wave data were used in this study to obtain more stable estimates. The simulation used a 2 (number of individuals: 210 or 390) × 3 (number of within-subject covariance structures: AR(1), UN(1), or ID) factorial design to generate the data. Because the study with the first scenario showed the low convergence rate (e.g., 54.10%) for the AR(1) covariance structure, a total of 4,000

replications for the AR(1) structure were generated. The other two covariance structures (i.e., ID and UN(1)) generated 2,000 replications for each condition using the Mplus6.11 Monte Carlo procedure (Muthén & Muthén, 1998-2010), yielding 14,000 total datasets. Each dataset was then analyzed using the four start models along with a parameter addition/deletion method. Correct model recovery was evaluated by several types of model selection indices/criteria as with the study of the first scenario.

3.2.1.1 Research design and model parameterization

The same numbers of participants (i.e., 210 for small and 390 for large sample size) were used for data generation. A quadratic growth trajectory was used as the true/target model. Similar to the scenario 1, the effect size equation [28] was used to select the population parameters. Given the values of δ and τ_{11}, I used 0.16 for both linear (α_{21}) and quadratic (α_{31}) growth coefficients, as used for the linear growth coefficient in scenario 1.

In the second scenario, three types of error variance-covariance structures were considered, which were ID, AR(1), and UN(1) structures. For the ID and AR(1) structures, σ^2 was set to be 1.00 and ρ was 0.80. For the UN(1) error structure, all the covariances were set to zero. The residual variance of the first time point was set to 1.00 and the following residual variances were set to be the power function of ρ (i.e., $\sigma_1^2 = 1.000$, $\sigma_2^2 = 0.800$, $\sigma_3^2 = 0.640$, $\sigma_4^2 = 0.512$, $\sigma_5^2 = 0.410$, $\sigma_6^2 = 0.328$, $\sigma_7^2 = 0.262$, $\sigma_8^2 = 0.210$) assuming that the reliability of the measurement increases over time (Kwok et al., 2007; Grimm & Widaman, 2010). Since there were fixed effects associated with the

quadratic growth rate, I added more variance covariance parameters of the between-subject variance (Φ) matrix. The elements in the Φ matrix were set to:

$$\Phi_{Quadratic} = \begin{bmatrix} .20 & .05 & .05 \\ .05 & .10 & .035 \\ .05 & .035 & .10 \end{bmatrix},$$

having the correlations (i.e., $r = \frac{\varphi_{xy}}{\sqrt{\varphi_{xx}\varphi_{yy}}}, x \neq y$) of 0.35 for the elements in the Φ matrix (Kwok et al., 2007). The same specification search strategies with the scenario 1 were adopted in this study: four start models, a parameter addition/deletion method, and model selection tools.

3.2.2 Results for scenario 2

In the second scenario, I examined the performance of a parameter addition/deletion method combined with the four different start models to recover the true population mean structure. As was done with the first study, the primary dependent variable was the rate of the correct model recovery, which was coded as a binary variable (i.e., 0 for fail and 1 for success). Percentage of replications with the correct model recovery was found using the same six model selection indices (i.e., LRT, ΔCFI, ΔRMSEA, ΔSRMR, ΔAIC, and ΔBIC).

3.2.2.1 Convergence rate

Table 5 presents the convergence rate of the correctly specified model for six conditions of population models (i.e., 2 sample sizes and 3 within-subject covariance structures). Given that the convergence rate for the AR(1) covariance structure was low from the previous study, the number of replications for the AR(1) structure model was increased to 4,000. Among the 2,297 (57.43%) and 2,454 (61.35%) converged results for

Table 5. Percentage of convergence for the true quadratic growth model for study 1 scenario 2

Mean	Cov	wave	*n*	Final #of successful replications[b]	%of replications
Quadratic	AR(1)	8	210	2297	57.43%
Quadratic	AR(1)	8	390	2454	61.35%
Quadratic	UN(1)	8	210	2000	100.00%
Quadratic	UN(1)	8	390	2000	100.00%
Quadratic	ID	8	210	2000	100.00%
Quadratic	ID	8	390	2000	100.00%

Note. [a]Mean: true quadratic growth model; [b]Number of replications for AR(1) = 4,000, Number of replications for UN(1) & ID = 2,000.

the two sample size conditions, the first 2,000 datasets for each sample size condition were selected for further model search study. All the replications for the true ID and UN(1) covariance structures were successfully converged. Therefore, all six conditions of population models used 2,000 datasets for further model specification search procedure.

Table 6 indicates the convergence rate for specifying the different starting points in terms of the two covariance structures (i.e., either ID or UN) and seven mean structures. As seen in the previous study, the results were not converged for higher-order polynomial model (i.e., 5^{th}-, and 6^{th}-order) when specifying the simplest ID residual structure because there were too many parameters estimated that should have been fixed to zero. Given that the true model was a quadratic growth trajectory, I constrained the between-subject variance-covariance elements for the 5^{th}- and 6^{th}-order terms to zero, which were located too far from the true growth trajectory. After doing that, most of the analyses were successfully converged as presented in Table 6.

Table 6. Percentage of convergence by specifying the various mean and covariance structure model for study 1 scenario 2

Wave	True mean	True cov[a]	Cov[b]	n	Mean[c]						
					0	1	2	3	4	5	6
8wave	Quadratic	AR(1)	id	210	100.00	100.00	100.00	99.75	99.25	99.70	99.45
8wave	Quadratic	AR(1)	id	390	100.00	100.00	100.00	99.95	99.10	99.65	99.15
8wave	Quadratic	AR(1)	un	210	100.00	100.00	100.00	100.00	100.00	100.00	99.95
8wave	Quadratic	AR(1)	un	390	100.00	100.00	100.00	100.00	100.00	100.00	99.95
8wave	Quadratic	ID	id	210	100.00	100.00	100.00	100.00	99.95	99.95	100.00
8wave	Quadratic	ID	id	390	100.00	100.00	100.00	100.00	100.00	100.00	99.95
8wave	Quadratic	ID	un	210	100.00	100.00	100.00	100.00	100.00	100.00	100.00
8wave	Quadratic	ID	un	390	100.00	100.00	100.00	100.00	100.00	100.00	100.00
8wave	Quadratic	UN(1)	id	210	100.00	100.00	100.00	100.00	99.85	99.90	99.90
8wave	Quadratic	UN(1)	id	390	100.00	100.00	100.00	100.00	99.80	99.95	99.90
8wave	Quadratic	UN(1)	un	210	100.00	100.00	100.00	100.00	100.00	100.00	100.00
8wave	Quadratic	UN(1)	un	390	100.00	100.00	100.00	100.00	100.00	100.00	99.95

Note. [a]True cov: true covariance structure (i.e., AR(1), ID, or UN(1)); [b]Cov: specified covariance structure (i.e., id=Identity, un=Unstructured); [c]Mean: the degree of polynomial growth term (e.g., 0 = intercept-only model, 1 = linear growth model, 2 = quadratic growth model, etc.).

3.2.2.2 True model recovery

Table 7 presents the hit rate of the model selection criteria by different start models and search methods for each simulation condition. The first column indicates the starting point for the model specification search. As the same with the first scenario, there were four starting points in terms of the covariance and mean structures (see Figure 4): (1) simplest covariance with simplest mean structure, (2) simplest covariance with most complex/saturated mean structure, (3) most complex/saturated covariance with simplest mean structure, and (4) most complex/saturated covariance and mean structure.

Simplest covariance with simplest mean structure. In Table 7, the first six rows, labeled Start (1), present the hit rate of the model selection criteria when the simplest covariance and the mean structures were specified. There were three conditions related to the true covariance structure (i.e., ID, UN(1), and AR(1)). For the true ID structure condition, LRT, ΔSRMR, ΔAIC, and ΔBIC showed successful performance with the average recovery rate of 94.53%, 98.40%, 98.40%, and 100.00%, respectively. In other words, except the two model selection indices, ΔCFI and ΔRMSEA, other searching tools were able to find the true mean model with specifying the true covariance structure. For the true UN(1) structure condition, however, the hit rates of LRT and ΔAIC were noticeably decreased to 16.79% and 28.12%. The hit rates of ΔSRMR and ΔBIC were also reduced to 84.87% and 85.47%, respectively. For the true AR(1) covariance structure, the hit rates of all model selection criteria went down to zero (or close to zero) except ΔSRMR, which had the hit rate of 92.38%. ΔCFI and ΔRMSEA showed zero percent of recovery rate for all simulation conditions. As presented in Table 8, the two fit

Table 7. Percentage of true model recovery by the model selection criteria for study 1 scenario 2[a]

Start	Cov[b]	Mean[c]	True Cov	*N*	LRT	ΔCFI	ΔRMSEA	ΔSRMR	ΔAIC	ΔBIC
(1)	ID	Step-up	ID	210	94.05	0.00	0.00	97.20	98.25	100.00
(1)	ID	Step-up	ID	390	95.00	0.00	0.00	99.60	98.55	100.00
(1)	ID	Step-up	UN(1)	210	28.63	0.00	0.00	81.93	44.89	93.59
(1)	ID	Step-up	UN(1)	390	4.95	0.00	0.00	87.80	11.35	77.35
(1)	ID	Step-up	AR(1)	210	0.25	0.00	0.00	91.35	0.25	2.70
(1)	ID	Step-up	AR(1)	390	0.05	0.00	0.00	93.40	0.05	0.05
(2)	ID	Top-down	ID	210	80.75	99.25	83.40	91.20	93.50	98.60
(2)	ID	Top-down	ID	390	81.35	100.00	90.50	98.20	94.00	99.20
(2)	ID	Top-down	UN(1)	210	0.25	9.16	62.96	73.97	0.90	43.79
(2)	ID	Top-down	UN(1)	390	0.00	3.20	65.50	84.50	0.05	3.00
(2)	ID	Top-down	AR(1)	210	0.25	3.20	3.95	87.00	0.25	1.75
(2)	ID	Top-down	AR(1)	390	0.00	0.20	0.50	91.60	0.05	0.05
(3)	UN	Step-up	ID	210	93.30	90.60	44.95	75.95	93.90	91.55
(3)	UN	Step-up	ID	390	95.60	82.35	48.00	84.85	98.80	99.35
(3)	UN	Step-up	UN(1)	210	93.99	54.95	48.45	91.14	95.45	94.59
(3)	UN	Step-up	UN(1)	390	93.85	64.50	55.50	96.65	97.95	99.25
(3)	UN	Step-up	AR(1)	210	92.65	8.10	36.10	72.65	91.35	87.50
(3)	UN	Step-up	AR(1)	390	94.40	3.30	37.15	78.65	98.10	98.75
(4)	UN	Top-down	ID	210	81.65	99.70	57.65	89.35	94.40	97.05
(4)	UN	Top-down	ID	390	81.40	100.00	63.10	97.30	94.10	97.90
(4)	UN	Top-down	UN(1)	210	80.18	99.75	58.46	94.29	94.09	97.05
(4)	UN	Top-down	UN(1)	390	79.65	100.00	63.15	99.10	93.90	98.00
(4)	UN	Top-down	AR(1)	210	81.80	94.80	60.10	99.00	94.25	96.95
(4)	UN	Top-down	AR(1)	390	80.70	99.10	60.40	99.95	94.30	98.25

Note. [a]True mean structure: quadratic growth model; [b]Start: start model: (1) the simplest covariance and mean structure, (2) the simplest covariance and the most complex mean structure, (3) the most complex covariance and the simplest mean structure, (4) the most complex covariance and mean structure; [c]Cov: specification of the covariance structure (i.e., ID = identity, UN = unstructured); [d]Mean: Step-up = parameter addition method starting from an intercept-only model (i.e., intercept-only model →linear model →quadratic model), Top-down = parameter deletion method starting from the most complex model (i.e., 6th order polynomial model →5th order polynomial model →4th order polynomial model →cubic model → quadratic model)

Table 8. Percentage of replications favoring a model with more parameters for study 1 scenario 2 (AR(1))[a]

Cov[b]	*n*	Mean[c]	LRT	ΔCFI	ΔRMSEA	ΔSRMR	ΔAIC	ΔBIC
ID	210	0 vs. 1	100.00	0.00	0.00	100.00	100.00	100.00
ID	210	1 vs. 2	100.00	100.00	100.00	100.00	100.00	100.00
ID	210	2 vs. 3	99.75	95.50	91.95	8.60	99.75	97.30
ID	210	3 vs. 4	92.45	22.90	43.45	4.35	84.70	19.20
ID	210	4 vs. 5	7.05	0.00	0.25	0.00	2.30	1.25
ID	210	5 vs. 6	4.80	0.00	0.10	0.00	1.25	0.45
ID	390	0 vs. 1	100.00	0.00	0.00	100.00	100.00	100.00
ID	390	1 vs. 2	100.00	100.00	100.00	100.00	100.00	100.00
ID	390	2 vs. 3	99.95	99.55	98.40	6.60	99.95	99.95
ID	390	3 vs. 4	98.95	20.15	43.95	1.80	98.30	65.00
ID	390	4 vs. 5	7.15	0.00	0.00	0.00	2.95	1.00
ID	390	5 vs. 6	5.25	0.00	0.00	0.00	1.10	0.40
UN	210	0 vs. 1	97.90	8.70	39.55	73.00	93.05	88.40
UN	210	1 vs. 2	100.00	94.80	100.00	100.00	100.00	100.00
UN	210	2 vs. 3	5.40	0.00	9.05	0.40	1.80	0.95
UN	210	3 vs. 4	4.10	0.00	10.60	0.20	0.90	0.45
UN	210	4 vs. 5	5.00	0.00	120.50	0.15	1.50	0.85
UN	210	5 vs. 6	5.15	0.00	16.25	0.25	1.60	0.85
UN	390	0 vs. 1	100.00	3.35	40.10	78.65	99.80	99.25
UN	390	1 vs. 2	100.00	99.10	100.00	100.00	100.00	100.00
UN	390	2 vs. 3	5.60	0.00	7.95	0.00	1.70	0.50
UN	390	3 vs. 4	4.75	0.00	9.45	0.00	1.40	0.45
UN	390	4 vs. 5	4.85	0.00	11.05	0.00	1.30	0.30
UN	390	5 vs. 6	5.70	0.00	16.30	0.05	1.40	0.55

Note. [a]True model: quadratic growth model with the AR(1) error structure; [b]Error variance-covariance structure: ID = identity, UN(1) = main diagonal banded, UN = unstructured; [c]Mean: the degree of polynomial growth term. (e.g., 0 = intercept-only model, 1 = linear growth model, 2 = quadratic growth model, etc.)

indices failed to favor the linear growth model over the intercept-only model because CFI for both intercept-only and linear growth models were 0.00, indicating that the model fit was bad; likewise, RMSEA for both intercept-only and linear models were 1.00, indicating the bad fit. For the simplest covariance and mean structure start model, only ΔSRMR performed successfully with the average hit rate of 91.88%. All other model selection indices performed inconsistently depending on the simulation condition. Although LRT, ΔAIC, and ΔBIC performed well for the true ID structure condition, it was not the ideal method given that the true covariance structure is unknown in the real data analysis.

Simplest covariance with most complex mean structure. The next six rows in Table 7, start model (2), indicate the recovery rates of the model selection indices for specifying the simplest covariance with the most complex mean structure. Similar to the pattern of the start model (1), most of the selection indices performed well when the true model was the ID covariance structure. However, when the true model had different covariance structures (i.e., UN(1) and AR(1)), the average hit rates of the LRT, ΔCFI, ΔRMSEA, ΔAIC, and ΔBIC for the two sample size conditions decreased to 0.13%, 1.70%, 2.23%, 0.15%, and 0.90%, respectively. As with the start model (1), ΔSRMR recovered the true model relatively well with the average hit rate of 89.30%.

Most complex covariance with simplest mean structure. To search for the true mean structure, the start model (3) in Table 7 used the most complex covariance structure (UN) with the simplest intercept-only model for the starting point. Regardless of the simulation conditions, LRT, ΔAIC, and ΔBIC recovered the true mean model

successfully with the average rate of 93.97%, 95.93%, and 95.17%, respectively. The overall hit rate of ΔSRMR across the simulation conditions was 83.32%. ΔSRMR, ΔAIC, and ΔBIC were more sensitive with the larger sample size showing the higher hit rate (i.e., 86.72%, 98.28%, and 99.12%) than the small sample size condition (i.e., 79.91%, 93.57%, and 91.21%). Although ΔCFI successfully recovered the true ID covariance structure by specifying the UN covariance structure with the average hit rate of 86.48% for the two sample size, it went down to 5.70% for the true AR(1) covariance structure condition. As shown in Table 8 to Table 10, most of the cases failed to favor the linear growth model over the intercept-only model. In Table 8, the ΔRMSEA hit rate ranged from 36.10% to 55.50% across the simulation conditions.

Most complex covariance with most complex mean structure. The last start model was for specifying both the most complex covariance and mean structure. Among the four start models, this start model (4) performed best with the overall hit rates of 98.89%, 97.53%, and 96.50% for ΔCFI, ΔBIC, and ΔSRMR, respectively, across the simulation conditions. ΔAIC also performed well, having an average hit rate of 94.17%. Regardless of the true covariance structure condition, those four model selection indices were able to recover the true mean trajectory successfully (i.e., more than 9 successes in 10 attempts). LRT showed the higher hit rate for the true UN(1) structure condition (i.e., 94.35%) than the other two covariance structures (i.e., 81.39%). Nevertheless, the overall LRT hit rate was still 85.71%, which was more than 4 successes in 5 attempts. ΔRMSEA showed the lowest performance to recover the true mean structure with the average hit rate of 60.48%.

Table 9. Percentage of replications favoring a model with more parameters for study 1 scenario 2 (ID)[a]

Cov[b]	*n*	Mean[c]	LRT	ΔCFI	ΔRMSEA	ΔSRMR	ΔAIC	ΔBIC
ID	210	0 vs. 1	100.00	0.00	0.00	99.80	100.00	100.00
ID	210	1 vs. 2	100.00	100.00	100.00	100.00	100.00	100.00
ID	210	2 vs. 3	5.95	0.25	5.35	2.65	1.75	0.00
ID	210	3 vs. 4	5.20	0.50	7.00	6.15	1.60	0.00
ID	210	4 vs. 5	5.10	0.00	2.60	0.00	1.70	0.80
ID	210	5 vs. 6	4.65	0.00	2.50	0.00	1.50	0.60
ID	390	0 vs. 1	100.00	0.00	0.00	100.00	100.00	100.00
ID	390	1 vs. 2	100.00	100.00	100.00	100.00	100.00	100.00
ID	390	2 vs. 3	5.00	0.00	2.85	0.40	1.45	0.00
ID	390	3 vs. 4	4.20	0.00	4.20	1.40	1.20	0.00
ID	390	4 vs. 5	5.35	0.00	1.40	0.00	1.65	0.30
ID	390	5 vs. 6	5.55	0.00	1.20	0.00	1.80	0.50
UN	210	0 vs. 1	98.00	70.70	49.60	76.40	95.15	92.05
UN	210	1 vs. 2	100.00	99.90	100.00	100.00	100.00	100.00
UN	210	2 vs. 3	4.85	0.05	10.00	0.55	1.30	0.55
UN	210	3 vs. 4	5.10	0.05	10.90	1.25	1.50	0.75
UN	210	4 vs. 5	5.15	0.05	13.55	2.95	1.60	0.90
UN	210	5 vs. 6	5.00	0.05	16.50	6.10	1.50	0.80
UN	390	0 vs. 1	100.00	82.35	51.90	84.90	99.95	99.90
UN	390	1 vs. 2	100.00	100.00	100.00	100.00	100.00	100.00
UN	390	2 vs. 3	4.40	0.00	7.20	0.05	1.15	0.55
UN	390	3 vs. 4	4.40	0.00	8.05	0.15	1.30	0.50
UN	390	4 vs. 5	5.30	0.00	11.15	0.55	1.75	0.40
UN	390	5 vs. 6	5.95	0.00	15.10	1.95	1.75	0.65

Note. [a]True model: quadratic growth model with the ID error structure; [b]Error variance-covariance structure: ID = identity, UN(1) = main diagonal banded, UN = unstructured; [c]Mean: the degree of polynomial growth term. (e.g., 0 = intercept-only model, 1 = linear growth model, 2 = quadratic growth model, etc.)

Table 10. Percentage of replications favoring a model with more parameters for study 1 scenario2 (UN)[a]

Cov[b]	*n*	Mean[c]	LRT	ΔCFI	ΔRMSEA	ΔSRMR	ΔAIC	ΔBIC
ID	210	0 vs. 1	100.00	0.00	0.00	99.85	100.00	100.00
ID	210	1 vs. 2	100.00	100.00	100.00	100.00	100.00	100.00
ID	210	2 vs. 3	71.31	20.97	0.65	18.02	55.11	6.41
ID	210	3 vs. 4	99.30	89.29	36.74	8.66	97.95	53.40
ID	210	4 vs. 5	5.06	0.00	0.00	0.00	1.55	0.50
ID	210	5 vs. 6	4.70	0.00	0.05	0.00	1.55	0.70
ID	390	0 vs. 1	100.00	0.00	0.00	99.95	100.00	100.00
ID	390	1 vs. 2	100.00	100.00	100.00	100.00	100.00	100.00
ID	390	2 vs. 3	95.05	13.45	0.00	12.15	88.65	22.65
ID	390	3 vs. 4	99.80	96.35	34.50	3.35	99.80	95.80
ID	390	4 vs. 5	5.95	0.00	0.00	0.00	1.80	0.45
ID	390	5 vs. 6	5.80	0.00	0.00	0.00	1.15	0.35
UN	210	0 vs. 1	99.15	55.16	54.30	91.34	97.30	95.60
UN	210	1 vs. 2	100.00	99.75	100.00	100.00	100.00	100.00
UN	210	2 vs. 3	5.16	0.00	10.51	0.20	1.85	1.00
UN	210	3 vs. 4	5.26	0.00	10.86	0.55	1.35	0.70
UN	210	4 vs. 5	5.11	0.00	12.46	1.30	1.30	0.65
UN	210	5 vs. 6	5.51	0.00	15.42	3.65	1.40	0.60
UN	390	0 vs. 1	100.00	64.50	60.45	96.65	100.00	99.95
UN	390	1 vs. 2	100.00	100.00	100.00	100.00	100.00	100.00
UN	390	2 vs. 3	6.15	0.00	8.60	0.00	2.05	0.70
UN	390	3 vs. 4	4.85	0.00	8.50	0.15	1.35	0.55
UN	390	4 vs. 5	5.20	0.00	11.00	0.20	1.55	0.45
UN	390	5 vs. 6	5.65	0.00	14.70	0.55	1.35	0.40

Note. [a]True model: quadratic growth model with the UN(1) error structure; [b]Error variance-covariance structure: ID = identity, UN(1) = main diagonal banded, UN = unstructured; [c]Mean: the degree of polynomial growth term. (e.g., 0 = intercept-only model, 1 = linear growth model, 2 = quadratic growth model, etc.)

4. STUDY 2: MODEL SPECIFICATION SEARCH FOR THE OPTIMAL COVARIANCE STRUCTURE IN LATENT GROWTH MODELING (LGM): A MONTE CARLO STUDY

In the previous study, I limited the model specification search to find the optimal mean trajectory (i.e., linear growth for the scenario 1; quadratic growth for the scenario 2) for simulated data. In this study, the model search was continued to find the true/target covariance structure for the datasets that recovered the true mean structure. Although the impact of model misspecification on the within-subject (error) variance-covariance structure is substantial (Ferron et al., 2002; Kwok et al., 2007; Sivo et al., 2005), it has rarely been adapted to model building procedure in practice. In this study, I investigated the optimal search strategy for finding the correct/target error variance-covariance structure.

4.1 Method

Search was limited to the datasets that successfully recovered the true mean structure in the previous study. Given that the best strategy to find the target mean structure was saturating both the covariance and mean structure, I used the datasets recovered by using the most complex covariance and mean structure (i.e., model (4) in Figure 4) as the start model. As with Study 1, there were 10 types of the true growth models, which were comprised of two number of measurement waves (4 or 8), two sample sizes (210 or 390), and three types of within-subject variance-covariance

structures (ID, UN(1), or AR(1)). The number of datasets for each true model adopted for this study was presented in Table 11.

4.1.1 Stationary vs. Nonstationary

To continue the model search for the true covariance structure, I adopted the two patterns of the error variance-covariance structures, which were categorized by the stationarity of the residual variances. An error structure can be considered as stationary if the residual variances are invariant and the equal covariances at lag k (i.e., k = T – 1; T is the number of measurements) are hold. The stationary error covariance structure includes the AR(1), ARMA(1,1), CS, and TOEP structure, which have the equal variance and equal covariances at lag k. For example, AR(1) structure in matrix [14] shows the equal variances across the repeated measures and auto-correlation ρ for adjacent observations. Given that the AR(1) and ARMA(1,1) were commonly used stationary residual structures for equally spaced measurements, these two structures were adopted for this study. Therefore, for the stationary covariance structures, the model comparisons were proceeded through the ID, AR(1), ARMA(1,1) and UN structures. On the contrary, when the residual variances are variant (not equal) across the repeated measurements, the error covariance structure is considered as nonstationary. The nonstationary covariance structures include the heterogeneous type of covariance structures (e.g., heterogeneous TOEP, heterogeneous CS) and the unstructured with q bands (i.e., q = 1,…,T). In this study, for the nonstationary covariance structures, I used the UN(1), which specifies a completely general (unstructured) covariance matrix for the

Table 11. The number and percentage of datasets recovering the true mean structure by saturating the mean and covariance structures

Wave	True cov	True mean	*n*	LRT	ΔCFI	ΔRMSEA	ΔSRMR	ΔAIC	ΔBIC	Percent converged
4	AR(1)	Linear	210	1888	1961	1677	1729	1961	1972	98.29
4	AR(1)	Linear	390	1887	1997	1696	1871	1966	1987	99.02
8	AR(1)	Linear	210	1542	2000	1046	1834	1849	1934	100.00
8	AR(1)	Linear	390	1541	2000	1175	1950	1852	1945	100.00
8	AR(1)	Quadratic	210	1636	1896	1202	1980	1885	1939	96.58
8	AR(1)	Quadratic	390	1614	1982	1208	1999	1886	1965	97.47
8	ID	Quadratic	210	1633	1994	1153	1787	1888	1941	92.34
8	ID	Quadratic	390	1628	2000	1262	1946	1882	1958	92.33
8	UN(1)	Quadratic	210	1602	1993	1168	1884	1880	1939	90.92
8	UN(1)	Quadratic	390	1593	2000	1263	1982	1878	1960	90.44

first band of matrix. Therefore, the model comparisons were proceeded through the ID, UN(1), and UN structures.

4.1.2 Start models and specification search method

After setting up the corresponding true mean structure, within-subject covariance structure (θ_δ) for a population model was searched. There were two start models in terms of the covariance structures, the simplest ID structure and the most complex UN structure. When using the step-up approach, the simplest covariance structure was used as the starting model (i.e., ID structure) which was then compared to the other more complex covariance structure model (e.g., AR(1)) with the use of different model comparison methods. Searching was stopped when the result was not statistically significant (i.e., model with simpler covariance structure was selected as the best fitting mean growth). With respect to the between-subject variance structure, unstructured (UN) structure (i.e., estimating all the elements in Φ matrix) was always used as the default.

When using the top-down approach, the most complex covariance structure was used as the starting model (i.e., UN structure) which was then compared to the other simpler covariance structure model (e.g., UN(1)). Given that the UN structure estimates all the elements in θ_δ matrix, the degrees of freedom for the covariance structure is zero, which leads to constrain all the elements in Φ matrix to zero. If the result of model comparison was statistically significant, searching was stopped and the more complex covariance structure was selected as the best fitting model. Again, each model searching method (i.e., LRT, information criteria, ΔGFI) was used to reach the true/target model and the hit rate would be summarized. For the true quadratic model with the UN(1)

covariance structure, the step-up or top-down method was not considered because UN(1) was located in the middle of the ID and UN structures.

4.1.3 Evaluation criteria and specification search procedure

As with Study 1, the same model comparison methods (i.e., LRT, ΔCFI, ΔRMSEA, ΔSRMR, ΔAIC, and ΔBIC) were used in searching for the true within-subject covariance structure. LRT was used since the residual structures along with the stationary or nonstationay types of structures are nested with each other. Regardless of nesting structure, ΔGFI (i.e., $GFI_c - GFI_{uc}$) and Δinformation criteria (i.e., ΔAIC and ΔBIC) were also used for the model searching procedure. The same cutoff criteria (i.e., Chen, 2007: 0.01 for ΔCFI, 0.015 for ΔRMSEA, 0.01; Burnham & Anderson, 1998: 4 for ΔAIC; Raftery,1996: 2 for ΔBIC) were used in this study. For example, if the ΔCFI was at least 0.01, the model comparison was statistically significant and searching was stopped. A model with more parameters (i.e., a model with more complex covariance structure) was selected for the best fitting model.

4.2 Results

In Study 2, I examined the performance of the model comparison methods with the step-up and top-down approaches to recover the true population covariance structure. As was done with Study 1, the primary dependent variable was the rate of the correct model recovery. The hit rates of LRT, ΔCFI, ΔRMSEA, ΔSRMR, ΔAIC, and ΔBIC were summarized.

4.2.1 Convergence rate

The last column in Table 11 presents the average convergence rate for each true mean structure model with specifying the different types of covariance structures (i.e., ID, AR(1), ARMA(1,1), UN(1), and UN). The average convergence rate across 10 models was 95.74%. Models with true quadratic growth, UN(1) structure, 390 cases, and 8 waves had the lowest mean convergence rate (mean = 90.44%). Models with true linear growth, AR(1) structure, and 8 waves had the highest mean convergence rate (mean = 100.00%) for both sample size conditions.

4.2.2 True model recovery

Table 12 presents the hit rate of model selection criteria reaching the true within-subject covariance structure. The hit rate indicates the correct model recovery rate from the corresponding number of datasets, which are presented in Table 11. The first column indicates the stationarity of the error variance structures for the search procedure (i.e., S for stationary and NS for nonstationary covariance structures). For the stationary covariance structures, ID vs. AR(1), AR(1) vs. ARMA(1,1), and ARMA(1,1) vs. UN structure were compared to test the model fit improvement. Step-up method started the model comparison from the ID structure to more complex structure, whereas top-down method started from the UN structure. For the nonstationary covariance structures, ID vs. UN(1) and UN(1) vs. UN were compared. Given that two of the three true covariance structures (e.g., ID and AR(1)) were close to simplest covariance structure (i.e., ID), some of the model selection criteria performed better when using the step-up method than the top-down method.

Table 12. Percentage of true covariance structure recovery by the model selection criteria for study 2

Stationary[a]	Method[b]	Wave	TRUE mean	TRUE cov	*N*	LRT	ΔCFI	ΔRMSEA	ΔSRMR	ΔAIC	ΔBIC
S	Step-up	4	Linear	AR(1)	210	94.86	86.08	93.02	59.69	94.90	92.80
S	Step-up	4	Linear	AR(1)	390	96.61	97.20	95.93	83.54	99.03	99.60
S	Step-up	8	Linear	AR(1)	210	94.42	93.90	98.66	97.00	98.38	99.02
S	Step-up	8	Linear	AR(1)	390	94.81	99.00	99.66	99.95	98.11	99.38
S	Step-up	8	Quadratic	AR(1)	210	96.58	99.95	99.33	1.92	98.73	99.43
S	Step-up	8	Quadratic	AR(1)	390	96.84	100.00	99.75	9.60	98.62	99.49
S	Step-up	8	Quadratic	ID	210	93.63	100.00	98.35	100.00	97.88	99.12
S	Step-up	8	Quadratic	ID	390	95.15	100.00	99.29	100.00	98.62	99.39
S	Top-down	4	Linear	AR(1)	210	90.68	83.89	73.40	39.91	93.42	92.80
S	Top-down	4	Linear	AR(1)	390	92.05	96.85	78.66	66.22	97.66	99.60
S	Top-down	8	Linear	AR(1)	210	89.82	92.45	71.99	3.49	98.32	99.12
S	Top-down	8	Linear	AR(1)	390	90.46	99.00	80.43	14.51	98.00	99.38
S	Top-down	8	Quadratic	AR(1)	210	91.56	99.79	79.45	1.62	98.73	99.43
S	Top-down	8	Quadratic	AR(1)	390	92.07	99.95	80.88	8.40	98.62	99.49
S	Top-down	8	Quadratic	ID	210	88.85	94.68	80.23	38.61	97.78	99.07
S	Top-down	8	Quadratic	ID	390	91.65	99.35	85.90	39.72	98.46	99.28
NS	Step-up	8	Quadratic	ID	210	94.67	99.25	93.06	80.13	98.89	100.00
NS	Step-up	8	Quadratic	ID	390	94.96	100.00	95.64	93.68	98.41	100.00
NS	Top-down	8	Quadratic	ID	210	88.92	94.13	65.83	1.34	98.57	100.00
NS	Top-down	8	Quadratic	ID	390	89.93	99.75	75.12	7.19	98.35	100.00
NS	-	8	Quadratic	UN(1)	210	94.51	98.44	75.94	5.57	99.95	100.00
NS	-	8	Quadratic	UN(1)	390	94.92	99.95	79.18	16.60	99.84	100.00

Note. [a]Type: S = stationary covariance structure (i.e., ID – AR(1) – ARMA(1,1) – UN), US = nonstationary covariance structure (i.e., ID – UN(1) – UN); [b]Method: Step-up = parameter addition method starting from the ID structure, Top-down = parameter deletion method starting from UN structure.

4.2.2.1 Likelihood ratio test hit rate

The average LRT hit rate across all the simulation conditions (i.e., number of measurements, true mean and covariance structures, and sample sizes) was 93.09% The range of the LRT hit rate was between 88.85% and 96.84%. LRT showed higher hit rate with the step-up method (i.e., average rate of 95.25%) than the top-down method (i.e., average rate of 90.06%). The average LRT hit rate for the search process through the stationary covariance structures was 93.13%. Within the stationary covariance structures, the step-up method recovered the true covariance structure with the average rate of 95.36% while the top-down method showed the hit rate of 90.89%. The results showed that the LRT hit rates with the large number of subjects (n = 390) were always higher than the hit rates with the small number of subjects (n = 210) irrespective of the searching methods. Table 13 shows the details of the model comparisons between the two adjacent covariance structures.

The average LRT hit rate by the nonstationary covariance structures across the conditions was 92.99%. As same with using the stationary structures, the step-up method showed higher hit rate (mean = 94.82%) than the top-down method (mean = 89.43%). The details of the model comparison under the nontationary searching procedures are presented in Table 14.

4.2.2.2 ΔGFI hit rate

As following Chen's (2007) guidelines, ΔCFI, ΔRMSEA, and ΔSRMR were used for model comparison. ΔCFI showed the highest hit rate among the three fit indices. The average ΔCFI hit rate across all conditions was 96.98%, while the average

Table 13. Percentages of replications favoring more complex structure using the process of the stationary covariance structure

Wave	True cov	True mean	n	Comparison	LRT	ΔCFI	ΔRMSEA	ΔSRMR	ΔAIC	ΔBIC
4	AR(1)	Linear	210	ID - AR(1)	98.73	86.59	98.45	59.69	96.43	93.61
4	AR(1)	Linear	210	AR(1) - ARMA(1,1)	3.87	0.51	5.43	1.85	1.53	0.81
4	AR(1)	Linear	210	ARMA(1,1) vs UN	4.34	2.19	20.81	40.54	1.48	0.00
4	AR(1)	Linear	390	ID - AR(1)	100.00	97.20	100.00	83.54	100.00	99.90
4	AR(1)	Linear	390	AR(1) - ARMA(1,1)	3.39	0.00	4.07	0.11	0.97	0.30
4	AR(1)	Linear	390	ARMA(1,1) vs UN	4.98	0.35	17.81	22.34	1.48	0.00
8	AR(1)	Linear	210	ID - AR(1)	100.00	100.00	100.00	97.11	100.00	100.00
8	AR(1)	Linear	210	AR(1) - ARMA(1,1)	5.58	0.00	1.34	0.11	1.62	0.88
8	AR(1)	Linear	210	ARMA(1,1) vs UN	5.06	1.95	26.77	96.51	0.05	0.00
8	AR(1)	Linear	390	ID vs AR(1)	100.00	100.00	100.00	99.95	100.00	100.00
8	AR(1)	Linear	390	AR(1) vs ARMA(1,1)	5.19	0.00	0.34	0.00	1.89	0.62
8	AR(1)	Linear	390	ARMA(1,1) vs UN	4.61	0.00	19.23	85.49	0.11	0.00
8	AR(1)	Quadratic	210	ID vs AR(1)	99.94	99.95	100.00	1.92	99.95	99.95
8	AR(1)	Quadratic	210	AR(1) vs ARMA(1,1)	3.36	0.00	0.67	0.00	1.22	0.52
8	AR(1)	Quadratic	210	ARMA(1,1) vs UN	5.38	0.16	19.88	63.33	0.00	0.00
8	AR(1)	Quadratic	390	ID vs AR(1)	100.00	100.00	100.00	9.60	100.00	100.00
8	AR(1)	Quadratic	390	AR(1) vs ARMA(1,1)	3.16	0.00	0.25	0.00	1.38	0.51
8	AR(1)	Quadratic	390	ARMA(1,1) vs UN	4.96	0.05	18.87	41.57	0.00	0.00
8	ID	Quadratic	210	ID vs AR(1)	6.37	0.00	1.65	0.00	2.12	0.88
8	ID	Quadratic	210	AR(1) vs ARMA(1,1)	0.92	0.00	0.17	0.00	0.11	0.05
8	ID	Quadratic	210	ARMA(1,1) vs UN	4.41	5.32	18.56	61.39	0.00	0.00
8	ID	Quadratic	390	ID vs AR(1)	4.85	0.00	0.71	0.00	1.38	0.61
8	ID	Quadratic	390	AR(1) vs ARMA(1,1)	0.49	0.00	0.16	0.00	0.11	0.10
8	ID	Quadratic	390	ARMA(1,1) vs UN	3.44	0.65	13.39	60.28	0.11	0.00
8	UN(1)	Quadratic	210	ID vs AR(1)	6.93	0.00	0.00	0.00	2.10	1.24

Table 13 continued.

8	UN(1)	Quadratic	210	AR(1) vs ARMA(1,1)	0.12	0.00	0.00	0.00	0.00	0.00
8	UN(1)	Quadratic	210	ARMA(1,1) vs UN	54.87	54.99	54.20	54.83	55.05	6.45
8	UN(1)	Quadratic	390	ID vs AR(1)	6.47	0.00	0.00	0.00	1.65	0.56
8	UN(1)	Quadratic	390	AR(1) vs ARMA(1,1)	0.06	0.00	0.00	0.00	0.00	0.00
8	UN(1)	Quadratic	390	ARMA(1,1) vs UN	52.29	52.20	52.97	52.12	52.24	48.62

Table 14. Percentages of replications favoring more complex structure using the process of the nonstationary covariance structure

Wave	True cov	True mean	n	Comparison	LRT	ΔCFI	ΔRMSEA	ΔSRMR	ΔAIC	ΔBIC
4	AR(1)	Linear	210	ID - UN(1)	15.04	9.08	1.73	1.10	7.70	0.25
4	AR(1)	Linear	210	UN(1) - UN	95.60	86.89	97.14	99.19	87.56	40.77
4	AR(1)	Linear	390	ID - UN(1)	24.32	2.20	0.18	0.00	13.02	0.55
4	AR(1)	Linear	390	UN(1) - UN	99.89	97.15	99.82	99.95	99.80	91.55
8	AR(1)	Linear	210	ID - UN(1)	100.00	99.15	0.10	0.00	100.00	84.44
8	AR(1)	Linear	210	UN(1) - UN	100.00	100.00	100.00	100.00	100.00	100.00
8	AR(1)	Linear	390	ID - UN(1)	100.00	100.00	0.00	0.00	100.00	100.00
8	AR(1)	Linear	390	UN(1) - UN	100.00	100.00	100.00	100.00	100.00	100.00
8	AR(1)	Quadratic	210	ID - UN(1)	97.37	44.36	10.57	0.00	92.31	26.30
8	AR(1)	Quadratic	210	UN(1) - UN	100.00	100.00	99.67	94.90	100.00	3.35
8	AR(1)	Quadratic	390	ID - UN(1)	99.94	46.77	3.97	0.00	99.95	80.36
8	AR(1)	Quadratic	390	UN(1) - UN	100.00	100.00	100.00	97.35	100.00	90.69
8	ID	Quadratic	210	ID - UN(1)	5.33	0.75	6.94	19.87	1.11	0.00
8	ID	Quadratic	210	UN(1) - UN	6.00	5.12	28.01	98.10	0.32	0.00
8	ID	Quadratic	390	ID - UN(1)	5.04	0.00	4.36	6.32	1.59	0.00
8	ID	Quadratic	390	UN(1) - UN	5.28	0.25	20.68	91.73	0.05	0.00
8	UN(1)	Quadratic	210	ID - UN(1)	100.00	100.00	100.00	90.12	100.00	100.00
8	UN(1)	Quadratic	210	UN(1) - UN	5.49	1.56	24.06	94.21	0.05	0.00
8	UN(1)	Quadratic	390	ID - UN(1)	100.00	100.00	100.00	9.69	100.00	100.00
8	UN(1)	Quadratic	390	UN(1) - UN	5.08	0.05	20.82	83.35	0.16	0.00

Table 15. Final hit rates of true mean and covariance structure recovery

Stationary[a]	Method[b]	Wave	TRUE mean	TRUE cov	*n*	LRT	ΔCFI	ΔRMSEA	ΔSRMR	ΔAIC	ΔBIC
S	Step-up	4	Linear	AR(1)	210	89.55	84.40	78.00	51.60	93.05	91.50
S	Step-up	4	Linear	AR(1)	390	91.15	97.05	81.35	78.15	97.35	98.95
S	Step-up	8	Linear	AR(1)	210	72.80	93.90	51.60	88.95	90.95	95.85
S	Step-up	8	Linear	AR(1)	390	73.05	99.00	58.55	97.45	90.85	96.65
S	Step-up	8	Quadratic	AR(1)	210	79.00	94.75	59.70	1.90	93.05	96.40
S	Step-up	8	Quadratic	AR(1)	390	78.15	99.10	60.25	9.60	93.00	97.75
S	Step-up	8	Quadratic	ID	210	76.45	99.70	56.70	89.35	92.40	96.20
S	Step-up	8	Quadratic	ID	390	77.45	100.00	62.65	97.30	92.80	97.30
S	Top-down	4	Linear	AR(1)	210	85.60	82.25	61.55	34.50	91.60	91.50
S	Top-down	4	Linear	AR(1)	390	86.85	96.70	66.70	61.95	96.00	98.95
S	Top-down	8	Linear	AR(1)	210	69.25	92.45	37.65	3.20	90.90	95.85
S	Top-down	8	Linear	AR(1)	390	69.70	99.00	47.25	14.15	90.75	96.65
S	Top-down	8	Quadratic	AR(1)	210	74.90	94.60	47.75	1.60	93.05	96.40
S	Top-down	8	Quadratic	AR(1)	390	74.30	99.05	48.85	8.40	93.00	97.75
S	Top-down	8	Quadratic	ID	210	72.55	94.40	46.25	34.50	92.30	96.15
S	Top-down	8	Quadratic	ID	390	74.60	99.35	54.20	38.65	92.65	97.20
NS	Step-up	8	Quadratic	ID	210	77.30	98.95	53.65	71.60	93.35	97.05
NS	Step-up	8	Quadratic	ID	390	77.30	100.00	60.35	91.15	92.60	97.90
NS	Top-down	8	Quadratic	ID	210	72.60	93.85	37.95	1.20	93.05	97.05
NS	Top-down	8	Quadratic	ID	390	73.20	99.75	47.40	7.00	92.55	97.90
NS	-	8	Quadratic	UN(1)	210	75.70	98.10	44.35	5.65	98.00	98.90
NS	-	8	Quadratic	UN(1)	390	75.60	99.95	50.00	16.70	97.80	99.30

Note. [a]Type: S = stationary covariance structure (i.e., ID – AR(1) – ARMA(1,1) – UN), US = nonstationary covariance structure (i.e., ID – UN(1) – UN); [b]Method: Step-up = parameter addition method starting from the ID structure, Top-down = parameter deletion method starting from UN structure.

ΔRMSEA and ΔSRMR hit rates were 86.35% and 44.03%, respectively. The range of the ΔCFI hit rates was between 83.89% and 100.00%. Regardless of the searching methods, such as stationary process of the covariance structures and parameter addition/deletion methods, ΔCFI performed successfully to recover the true covariance structure with the average hit rate above 90%. For the small number of subjects and measurement waves (4-wave with 210 subjects), the average ΔCFI hit rate was 84.99%.

Within the search by stationary covariance structures, ΔRMSEA hit rate was 98.00% for the step-up method and 78.87% for the top-down method. Similarly, the nonstationary covariance structures revealed the higher hit rate with the step-up method (mean = 94.35%) than with the top-down method (mean = 70.48%). Although ΔSRMR showed similar pattern with other fit indices as favoring the step-up method regardless of the stationarity for the residual variance, ΔSRMR was sensitive to the simulation conditions. The ΔSRMR hit rate was ranged from 1.34% to 100.00% by simulation condition. For the true quadratic growth model with the ID structure for the small sample size, the step-up method using the stationary procedure showed 100.00% of hit rate, while the top-down method using the nonstationary procedure showed 1.34% of hit rate. Table 14 presents the details of the model comparisons between the two adjacent covariance structures under the specification of the nonstationary covariance structures. As shown in the table, for the condition above (i.e., true quadratic growth with ID for small sample size), 98.10% of replications failed to favor the UN(1) structure over the UN structure by ΔSRMR, which leaded the low recovery of the true ID structure.

4.2.2.3 ΔInformation criteria hit rate

As shown in Table 12, ΔAIC and ΔBIC showed the best performance along with ΔCFI in searching for the true within-subject covariance structure. The average ΔAIC and ΔBIC hit rates across all simulation conditions were 98.15% and 98.93%, respectively. The three fit indices performed almost perfectly regardless of the parameter addition/deletion methods. The standard deviations of the ΔAIC and ΔBIC hit rates suggested that the two selection criteria are very stable as well as successful. Surprisingly, ΔBIC showed 100.00% of hit rate to recover the true ID and UN(1) covariance structures when using the nonstationary process in searching for the correct covariance structure. In addition, except the condition for the small number of measurement waves and small sample size (mean = 92.80%), all the ΔBIC hit rates were above 99.00%.

Table 15 presents the final hit rate of the true model in terms of both mean and residual structures by searching the mean structure using the most complex/saturated mean and covariance structure. The percentage of each model selection criteria indicates the correct model recovery rate from the total number of 2,000 replications. As presented in the table, ΔBIC performed best to recover the true model (mean = 96.62%) followed by ΔCFI (mean = 96.44%) and ΔAIC (mean = 92.85%). Regardless of the search method (i.e., step-up or top-down) to find the optimal covariance structure, these three fit indices recovered the true model successfully (i.e., more than 9 successes in 10 attempts).

5. DISCUSSIONS AND CONCLUSIONS

5.1 Discussions

In this study, I examined the effectiveness of the model specification search by four different start models to recover the true growth model. To observe the growth change for a large scale longitudinal data, as Singer and Willett (2003) recommended, it is appropriate to build a mean growth trajectory first to summarize the individual's growth change. Then, it is also necessary to look into the inter-individual differences/variation and the intra-individual differences on the average growth rate to see how they differ. In a similar manner, the first part of the current study was conducted to search for the true mean structure using the six model selection tools, including LRT, ΔGFI, and Δinformation criteria. Further searching for the within-subject covariance structure was continued for the datasets that successfully recovered the true mean structure in the previous step.

5.1.1 Searching for the true mean structure

Findings from Study 1 suggested using the most complex covariance structure as the start model to search the optimal mean structure. In the first scenario of Study 1, the true models were the linear growth model with the unstructured between-subject covariance structure with the AR(1) error structure. For the true linear model with 4-wave data, the hit rates of the model selection indices for the two possible start models (i.e., either the simplest covariance structure or the most complex covariance structure) were compared. As shown in Figure 5 and Figure 8, saturating the within-covariance

structure performed substantively better than the simplest covariance structure to search the true mean trajectory by all the model selection criteria. Likewise, for the true linear growth with 8-wave data, specifying the covariance structure to be saturated showed high recovery rate of the true mean structure. Figure 6 and 7 present the performance of the simplest covariance structure, while Figure 9 and 10 show the performance of the most complex covariance structure. For example, under the same condition except the specification of the covariance structure, Figure 7 (i.e., the simplest covariance structure) shows almost 0% of hit rate of all fit indices on the graph, whereas Figure 10 shows a better than 90% recovery rate for ΔCFI, ΔSRMR, ΔAIC, and ΔBIC. Within specifying the most complex covariance structure, starting from the simplest mean structure (i.e., intercept-only model) performed better than the most complex mean structure (i.e., 6^{th} order polynomial model). Given that the true model (i.e., linear growth model) was closer to the simplest mean structure, the results seemed to favor the simplest mean structure start model. To determine whether the same results hold for different conditions, the next study was conducted.

In the second scenario, true population model was the quadratic growth model with more types of within-subject covariance structures (i.e., ID, AR(1), and UN(1)). For the single best fit index, ΔSRMR showed successful performance to search the optimal mean structure with the average hit rate of 90.60% regardless of the simulation conditions and the four different start models. However, ΔSRMR in the first scenario of Study 1, which set the linear growth model as the true mean structure, did not recover the true model successfully with specifying the simplest ID covariance structure (i.e.,

7.25%). Given that ΔSRMR was sensitive to the true mean structure, it was not recommended to use ΔSRMR solely in searching for the model specification.

On the other hand, both scenarios in Study 1 consistently revealed that saturating the covariance structure (i.e., starting from the most complex covariance structure) recovered the true mean model successfully by most of the model selection criteria. Although specifying the simplest ID structure recovered the true ID covariance structure model successfully (i.e., with the average hit rate above 90% for LRT, ΔSRMR, ΔAIC, and ΔBIC), the hit rates were substantially decreased when the true covariance model was not the ID structure. Despite of the high hit rate of some of the model selection indices, it is hard to argue that specifying the simplest covariance structure is an ideal method because the true covariance structure is unknown in the real data analysis. The correct model recovery needs to be robust to the simulation condition, such as, the true mean and covariance structures, sample sizes, and number of measurement waves. In terms of setting the mean structure starting point, except LRT and ΔAIC, which showed the highest hit rate with specifying the simplest mean structure model, other four indices favored specifying the start model for the mean structure to be saturated. In addition, the variations on the hit rates of the selection criteria were smaller with the most complex mean structure model.

In sum, specifying the saturated covariance structure with the most complex mean structure is considered as the best approach to search for the population mean trajectory. These results support the findings by Wu and West (2010) that saturating the covariance structure can detect the misspecification in the mean structure in latent

growth models. In their study, commonly used SEM fit indices (i.e., LRT, CFI, TLI, RMSEA, and SRMR) were sensitive to detect the model misspecification in the mean structure when the covariance matrix was saturated. This current study extended their findings to apply for the model searching procedure in LGM. In contrast, Ryoo (2011) used the simplest covariance structure (i.e., ID) in searching for the true mean structure in his simulation study within the linear mixed modeling framework. For simplicity, the specification search for the error variance structure was not included as one of the procedures to build an optimal growth model. However, findings from my study showed that specifying the simplest covariance structure cannot reach the true mean structure in most cases regardless of the model selection tools, including LRT, which was used in Ryoo's study.

5.1.2 Searching for the true covariance structure

The second part of this study was conducted to find the optimal strategy to search for the true covariance structure in LGM. Given that the specification of the covariance structure was followed by searching for the mean structure, searching was limited to the datasets that correctly specified the true mean structure. Here, the selected datasets were based on the start model (4) in Figure 4 (i.e., the most complex covariance structure with the most complex mean structure) because the start model (4) performed best to recover the true mean structure. Under the true mean structure, there were two starting points to reach the correct covariance structure: the simplest covariance structure (step-up) and the most complex covariance structure (top-down). The searching procedure was conducted

using the two types of covariance structures categorized by the stationarity (i.e., stationary or nonstationary) of the residual variances over time.

Findings from this second study suggested using LRT, ΔCFI, ΔAIC, and ΔBIC to search for the population covariance structure. Regardless of the starting points and the stationarity of the residual variances, the four model selection tools successfully recovered the true covariance structure with the average hit rate of above 90%. Specifically, ΔBIC was the most powerful criterion to search for the correct covariance structure as well as the correct mean structure.

5.2 Limitations of this study and directions for future research

The current study considered models without any other predictors to explain the growth change in outcome variables. In practice, however, researchers are typically interested in the effect of time-varying or time-invariant predictors as well as the rate of growth change. As Wu and West (2010) pointed out, including significant predictors of change in the model may improve the model fit as explaining more variance on the outcome variables (Singer & Willett, 2003). Future research can be conducted to examine how adding predictors in a model will change the correct model recovery by different start models and selection tools. In addition, as mentioned earlier, the functions of the time transformation are various for growth modeling, including polynomial, piecewise, exponential, logarithmic, and so on. This study used one of the functions, which was the polynomial, because it is widely used and well known by applied researchers. For future research, other functions of the time transformation can be adopted to see whether the same findings are hold in searching for the mean and

covariance structures. Similarly, there are a number of different forms of the error covariance structures other than the ones that I considered in this study (e.g., ID, AR(1), ARMA(1,1)). For example, Toeplitz, compound symmetry, or heterogeneous moving average structures can exist in the population growth model.

With respect to the start mean model, I used the highest-order polynomial model for the saturated mean structure in this study. However, it may not be a desirable start model because it is more likely to have a convergence problem due to having too many parameters in a model. Likewise, the intercept-only model may not be a reasonable start model for some cases in which researchers expect to see a definite growth change. To reduce these issues, it may be appropriate to adopt the visual inspection by the nonparametric approach, which draws the individual growth trajectory on a graph (Singer and Willett, 2003). Although this study was focused on the parametric approach, which compared the competing models by statistical fit indices, nonparametric method can also assist to determine the start mean structure model.

When building a model, there are three components to consider for the optimal model specification: fixed effects (i.e., average growth trajectory), random effects (i.e., variation on the average growth), and random error (i.e., variation within the individual's growth change). In the current study, I primarily focused on the specification search for the fixed effects and random error. For the random effects part, which is the between-subject variance-covariance structure, the unstructured (UN) structure was used as the default setting. As noted in the method, when convergence problems were encountered, I constrained the variances and covariances associated with the higher order growth term

(e.g., 5^{th}-, 6^{th}-order polynomial term) to zero. The similar procedures were used by Ryoo (2011) in his simulation study to find the true mean structure model. Given that the model was getting too cumbersome with all the random effect elements, he reduced the number of free parameters in the random effects by removing the variances and covariances related with the higher-order growth term. When a model recovered the true/correct fixed effects (mean structure) but not the true random effects, the search was assumed to be successful and termed as an "approximate model". He showed the similarity between the true model and the approximate model by visual inspection; the graphs were almost identical. Therefore, I did not have as much concern about the specification of the between-subject covariance structure as much as the growth trajectories and the within-subject covariance structure. Further research may incorporate the specification of the between-subject covariance structure to find the best fitting model.

5.3 Conclusion

This dissertation investigated the optimal strategy for the model specification search in the latent growth modeling. Given that we do not know the true model in real data analysis, we have to build a hypothesized model. Although developing an initial model based on the theory from prior research is favored, sometimes researchers may need to specify the starting model in the absence of theory. In this study, the effectiveness of the start models in searching for the true population model was examined using the simulated data. There were four possible start models in terms of the mean and covariance structures: the simplest mean and covariance structure model, the simplest mean and the most complex covariance structure model, the most complex

mean and the simplest covariance structure model, and the most complex mean and covariance structure model. To the extent that the model searching procedure is an exploratory analysis, researchers need to use model comparison tools to improve the model to fit the data. I compared the six model selection criteria to recover the true model: Likelihood ratio test (LRT), ΔCFI, ΔRMSEA, ΔSRMR, ΔAIC, and ΔBIC.

In Study 1, I conducted the model search for the correct/target *mean* structure. The results showed that specifying the most complex covariance structure (UN) with the most complex mean structure presented the highest recovery with the average rate above 90% using the ΔCFI, ΔBIC, ΔAIC, and ΔSRMR. With respect to the start models for the covariance structure, it was obvious that the saturated covariance structure outperformed the simplest covariance structure to recover the true mean trajectory. In terms of the mean structure (i.e., step-up or top-down), the recovery rate seemed to be affected by the location of the true mean structure.

In Study 2, the model specification search for the true covariance structure was conducted. The results showed that LRT, ΔCFI, ΔAIC, and ΔBIC successfully searched the population covariance structure regardless of the starting points and the stationarity of the residual variances; the average recovery rates were above 90% for all four fit indices. Specifically, ΔBIC was the most powerful criterion to search for the correct covariance structure as well as the correct mean structure.

In conclusion, when researchers conduct a longitudinal data analysis using the latent growth modeling, it is necessary to consider the mean structure as well as the within-subject variance-covariance structure to build a model. In order to set the start

model to build the statistical model, they need to saturate the within-subject covariance structure.

REFERENCES

Bentler, P. M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin*, 107: 238–246.

Biesanz, J. C., Deeb-Sossa, N., Papadakis, A. A., Bollen, K. A., & Curran, P. J. (2004). The role of coding time in estimating and interpreting growth curve models. *Psychological Methods, 9*(1), 30.

Bollen, K. A. (1989). *Structural equations with latent variables.* Wiley New York.

Bosker, R., Snijders, T., & Guldemond, H. (2003). PINT (power in two-level designs): Estimating standard errors of regression coefficients in hierarchical linear models for power calculations, user's manual. *Unpublished Manuscript*

Browne, M. W. & Cudeck, R. (1993). Alternative ways of assessing model fit. In K. A.Bollen*, &* J.Scott Long (Eds.), *Testing structural equation models (pp. 1–9).* Newbury Park, CA*:* Sage*.*

Burnham, K, P. & Anderson, D. R. (1998). Model selection and inference: A practical information-theoretic approach. New York: Springer-Verlag.

Chen, F. F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Structural Equation Modeling, 14*(3), 464-504.

Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling: A Multidisciplinary Journal, 9*(2), 233-255.

Curran, P. J., Obeidat, K., & Losardo, D. (2010). Twelve frequently asked questions about growth curve modeling. *Journal of Cognition and Development, 11*(2), 121-136.

Curran, P. J. (2003). Have multilevel models been structural equation models all along? *Multivariate Behavioral Research, 38,* 529-569.

Duncan, T. E., Duncan, S. C., & Strycker, L. A. (1999). *An introduction to latent variable growth curve modeling: Concepts, issues, and applications.* Mahwah, NJ: Erlbaum.

Everitt, B. (1998). Analysis of longitudinal data. beyond MANOVA. *The British Journal of Psychiatry, 172*(1), 7.

Fan, X., & Sivo, S. A. (2009). Using ΔGoodness-of-fit indexes in assessing mean structure invariance. *Structural Equation Modeling: A Multidisciplinary Journal, 16*(1), 54-69.

Ferrer, E., Hamagami, F., & McArdle, J. J. (2004). Modeling latent growth curves with incomplete data using different types of structural equation modeling and multilevel software. *Structural Equation Modeling: A Multidisciplinary Journal, 11*(3), 452-483.

Ferron, J., Dailey, R., & Yi, Q. (2002). Effects of misspecifying the first-level error structure in two-level models of change. *Multivariate Behavioral Research, 37*(3), 379-403.

Goldstein, H. (1995). *Multilevel statistical models* (2nd ed.). New York, NY: Wiley & Sons.

Green, S. B., Thompson, M. S., & Poirier, J. (1999). Exploratory analyses to improve model fit: Errors due to misspecification and a strategy to reduce their occurrence. *Structural Equation Modeling: A Multidisciplinary Journal, 6*(1), 113-126.

Grimm, K. J., & Widaman, K. F. (2010). Residual structures in latent growth curve modeling. *Structural Equation Modeling: A Multidisciplinary Journal, 17*(3), 424-442.

Hamaker, E. L., Dolan, C. V., & Molenaar, P. C. M. (2002). On the nature of SEM estimates of ARMA parameters. *Structural Equation Modeling, 9*(3), 347-368.

Hox, J. J. (2002). *Multilevel analysis techniques and applications* Mahwah, NJ: Lawrence Erlbaum Associates.

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling, 6*(1), 1-55.

Jöreskog, K. G., & Sörbom, D. (1981). *LISREL VI: Analysis of linear structural relationships by the method of maximum likelihood, instrumental variables, and least squares methods* Scientific Software.

Keselman, H., Algina, J., Kowalchuk, R. K., & Wolfinger, R. D. (1998). A comparison of two approaches for selecting covariance structures in the analysis of repeated measurements. *Communications in Statistics-Simulation and Computation, 27*(3), 591-604.

Khoo, S. T., West, S. G., Wu, W., & Kwok, O. M. (2006). Longitudinal methods. *Handbook of Multimethod Measurement in Psychology,* , 301–317.

Kline, R. B. *(*1998). *Principles and practice of structural equation modeling.* NY : The Guilford Press.

Kwok, O., West, S. G., & Green, S. B. (2007). The impact of misspecifying the within-subject covariance structure in multiwave longitudinal multilevel models: A monte carlo study. *Multivariate Behavioral Research, 42*(3), 557-592.

Kwok, O. M., Luo, W., & West, S. G. (2010). Using modification indexes to detect turning points in longitudinal data: A monte carlo study. *Structural Equation Modeling: A Multidisciplinary Journal, 17*(2), 216-240.

Leamer, E. E., & Leamer, E. E. (1978). *Specification searches: Ad hoc inference with nonexperimental data*. Boston: John Wiley.

Leite, W. L. & Stapleton, L. M. (2011). Detecting growth shape misspecifications in latent growth models: an evaluation of fit indexes, *The Journal of Experimental Education, 79*(4), 361-381.

Littell, R. C., Pendergast, J., & Natarajan, R. (2000). Tutorial in biostatistics: Modelling covariance structure in the analysis of repeated measures data. *Statistics in Medicine, 19*(1793), 1819.

Long, J. S. (1983). *Covariance structure models: An introduction to LISREL.* Beverly Hills, CA: Sage.

MacCallum, R. C., Kim, C., Malarkey, W. B., & Kiecolt-Glaser, J. K. (1997). Studying multivariate change using multilevel models and latent curve models. *Multivariate Behavioral Research*, *32*, 215-253.

MacCallum, R. (1986). Specification searches in covariance structure modeling. *Psychological Bulletin, 100*(1), 107-120.

Meredith, W., & Tisak, J. (1990). Latent curve analysis. *Psychometrika, 55*(1), 107-122.

Muthén, L. and Muthén, B. (1998–2011). *Mplus user's guide.* 6th ed., Los Angeles: Muthén & Muthén.

National Center for Education Studies. (1989). *NELS:88-Base year parent component, data file user's manual.* Washington, DC: U.S. Department of Education.

National Center for Education Studies. (2001). *ECLS-K first grade restricted-use user's manual.* Washington, DC: U.S. Department of Education.

Peugh, J. L., & Enders, C. K. (2010). Specification searches in multilevel structural equation modeling: A monte carlo investigation. *Structural Equation Modeling: A Multidisciplinary Journal, 17*(1), 42-65.

Preacher, K., Wichman, A., MacCallum, R., & Briggs, N. (2008). Latent growth curve modeling (vol. 157).Thousand Oaks, CA. Sage Publications, Inc.

Raftery, A. E. (1996). Bayesian model selection in social research. In P. V. Marsden (Ed.), Sociological Methodology (pp. 111-163). Oxford: Basil Blackwell.

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods* Sage Publications, Inc.

Raudenbush, S. W., & Liu, X. (2001). Effects of study duration, frequency of observation, and sample size on power in studies of group differences in polynomial change. *Psychological Methods, 6*(4), 387.

Ryoo, J. H. (2011): Model Selection with the Linear Mixed Model for Longitudinal Data, *Multivariate Behavioral Research, 46:4,* 598-624.

Sivo, S., Fan, X., & Witta, L. (2005). The biasing effects of unmodeled ARMA time series processes on latent growth curve model estimates. *Structural Equation Modeling: A Multidisciplinary Journal, 12*(2), 215-231.

Sivo, S. A., & Willson, V. L. (2000). Modeling causal error structures in longitudinal panel data: A monte carlo study. *Structural Equation Modeling, 7*(2), 174-205.

Snijders, T. A. B., & Bosker, R. J. (1993). Standard errors and sample sizes for two-level research. *Journal of Educational and Behavioral Statistics, 18*(3), 237.

Statistics Canada and Human Resources Development Canada. (1995) National Longitudinal Survey of Children: Overview of Survey Instruments for 1994–1995, Data Collection 1: *Statistics Canada Catalogue no. 95(2).* Minister of Industry Ontario, Canada.

Steiger, J. H. (1990). Structural model evaluation and modification: An interval estimation approach. *Multivariate Behavioral Research, 25*(2), 173-180.

Steiger, J. H., & Lind, J. C. (1980). Statistically based tests for the number of common factors. Paper presented at the *Annual Meeting of the Psychometric Society,* Iowa City, IA.

Suyapa, E., Silvia, M., & MacCallum, R. C. (1988). Some factors affecting the success of specification searches in covariance structure modeling. *Multivariate Behavioral Research, 23*(3), 297-326.

Widaman, K. F., & Thompson, J. S. (2003). On specifying the null model for incremental fit indices in structural equation modeling. *Psychological Methods, 8*(1), 16-37.

Wu, W., & West, S. G. (2010). Sensitivity of fit indices to misspecification in growth curve models. *Multivariate Behavioral Research, 45*(3), 420-452.

Wu, W., West, S. G., & Taylor, A. B. (2009). Evaluating model fit for growth curve models: Integration of fit indices from SEM and MLM frameworks. *Psychological Methods, 14*(3), 183.

Yu, C. Y. (2002). Evaluating cutoff criteria of model fit indices for latent variable models with binary and continuous outcomes. *Unpublished doctoral dissertation*, University of California , Los Angeles.

VITA

Name: Minjung Kim

Address: Dept. of Educational Psychology
c/o Dr. Oi-man Kwok (advisor)
Texas A&M University
College Station, TX 77843-4225

Email Address: mjkim.epsy@gmail.com

Education: B.A., Hanyang University, South Korea, 2002
M.Ed., Educational Psychology, Texas A&M University, United States, 2008
Ph.D., Educational Psychology, Texas A&M University, United States, 2012